HELSINGIN KAUPPAKORKEAKOULU

Testing Accuracy of Beta Mapping-Based VaR Models In Calculation of Equity Portfolio VaR

Rahoituksen pro gradu –tutkielma
5.4.2002
Janne Ryhänen

Laskentatoimen laitoksen

laitosneuvoston kokouksessa 21 / 5 2002 hyväksytty

arvosanalla erinomainen 90 pistettä

KTT Jari Käppi KTT Vesa Puttonen

HELSINGIN KAUPPAKORKEAKOULU TIIVISTELMÄ
Rahoituksen pro gradu –tutkielma
Tekijä: Janne Ryhänen 5.4.2002

Testing Accuracy of Beta Mapping-Based VaR Models in Calculation of Equity Portfolio VaR

Tutkimuksen tavoitteet

Tutkimuksen päämääräisenä tavoitteena on selvittää kuinka täsmällisiä VaR –lukuja neljä erilaista beta mappausta soveltavaa VaR –mallia tuottavat osakesalkuille; jokainen tutkittavista malleista käyttää yleistä varianssi-kovarianssi –lähestymistapaa. Toisena tavoitteena on pyrkiä mahdollisimman laajaan käytännönläheisyyteen, minkä vuoksi sekä päivittäisiä että viikoittaisia VaR –lukuja lasketaan ja tutkitaan kahdella suosituimmalla VaR –luottamustasolla, eli 95% ja 99%, osakesalkkujen sisältäessä 30, 50 ja 100 osaketta.

Lähdeaineisto

Lähdeaineisto saadaan Datastream –tietokannasta. Salkut kootaan satunnaisesti S&P 500 –indeksin likvideistä osakkeista. Osakkeille kerätään tietoa kokonaistuottoindekseistä vuosilta 1991-2000 sekä markkina-arvoista testausperiodin alussa 1995. S&P 500 –indeksille sekä sen kahdeksalle eri sektori-indeksille haetaan tietoa tuottoindekseistä aikavälillä 1991-2000.

Tutkimusmenetelmät

Modernien tapojen mukaisesti, tuotot lasketaan tutkimuksessa geometrisesti. Perinteisen varianssi-kovarianssi VaR -mallin VaR –luvut lasketaan käyttämällä hyväksi osakesalkun osakkeiden tuottohistoriaan perustuvaa kovarianssimatriisia. Beta mappausta soveltavat mallit estimoivat matriisia markkina –ja sektori-indeksien beta –korjattujen varianssien ja kovarianssien avulla. Tuottosarjojen variansseja ja kovariansseja estimoidaan painottamalla menneitä havaintoja tasapuolisesti. Pienimmän neliösumman (PNS) menetelmää käytetään betojen estimoimiseen. VaR mallien parametrejä estimoidaan 250, 500 ja 1000 päivän pituisilla periodeilla. *t*-testillä testataan voidaanko tilastollisesti hyväksyä, että osakesalkun odotettu tuotto on nolla. Anderson-Darling –testiä käytetään osakesalkkujen tuottojen normaalijakaumaoletukselle. VaR –mallien hyvyyttä testataan todennäköisyysaste –testeillä.

Tulokset

Yksikään tutkituista VaR –malleista ei kykene täysin täyttämään täsmällisille VaR –luvuille asetettuja vaatimuksia. Selkeästi epäsoveliaita ovat mallien tuottamat 99% VaR –luvut, sillä toteutuneet osakesalkkutuotot ylittävät näitä lukuja tilastollisesti merkitsevästi enemmän kuin 1% kaikista kerroista; tämä johtunee osakesalkkutuottojakauman paksummista hännistä kuin mitä normaalijakauma implikoi. Tutkitut 95% VaR –luvut ovat tässä mielessä täsmällisiä, mutta niiden heikkous on VaR-ylitysten riippuvuus toisistaan. Yleisesti ottaen, beta mappausta soveltavat VaR –mallit, jotka huomioivat myös yrityskohtaiset riskit, ovat varsin tasaveroisia perinteisen varianssi-kovarianssi VaR –mallin kanssa. Yritysriskiä noteeraamattomat, pelkistetyt beta mappaus VaR –mallit vaikuttavat olevan kompetentteja vain osakesalkun ollessa suhteellisen suuri ja sisältäessä vähintään 100 osaketta. Mappaus markkinaindeksiin näyttää suotavammalta kuin mappaus sektori-indekseihin.

Avainsanat

Value at Risk (VaR), beta mappaus, testaus, testaaminen, portfolioteoria

HELSINKI SCHOOL OF ECONOMICS
Masters thesis in finance
Author: Janne Ryhänen

ABSTRACT

5.4.2002

Testing Accuracy of Beta Mapping-Based VaR Models in Calculation of Equity Portfolio VaR

Objectives of the study

The main objective of the study is find out how accurate VaR measures do the four different beta mapping-based VaR models produce for equity portfolios; each model applies the common variance-covariance approach. Another objective is to be as practical as possible. Thus, both daily and weekly VaR measures are studied at the two most popular VaR confidence levels of 95% and 99% with portfolios including 30, 50 and 100 equities.

Data

The study's data is collected from Datastream database. Equity portfolios are formed in a random manner from the liquid S&P 500 constituent equities. For equities, daily total return indices are collected from the database for years 1991-2000, as well as market capitalisations at the beginning of the test period 1995. In addition, daily return indices are collected for the S&P 500 index and for eight S&P 500 sector indices for 1991-2000.

Methodology

Following the modern practices, geometric returns are used throughout the study. VaR measures of the traditional variance-covariance VaR model are calculated by using the covariance matrix of the portfolio's equities' historical return series. The study's beta mapping-based VaR models use the beta-corrected variances and covariances of the market and sector indices in deriving an estimate for the covariance matrix. Variances and covariances of the equity and index return series are estimated by equally weighting past observations. The ordinary least squares (OLS) procedures are used in beta estimation. The parameters of the VaR models are estimated applying an estimation period length of 250, 500 and 1000 days. The *t*-tests are used to test zero expected portfolio returns. The Anderson-Darling tests are applied to test normality of the portfolio returns. Furthermore, likelihood ratio tests are carried out in examining the accuracy abilities of the VaR models.

Results

None of the studied VaR models fulfils the requirements for highly accurate VaR measures. The VaR measures turn out to be really poor at the 99% VaR confidence level, since the VaR measure exceptions occur statistically significantly more than 1% of times; this seems to be due to equity portfolio distributions' fatter tails than implied by the normal distribution. The studied 95% VaR measures are on this respect accurate, but their exceptions fail to be independent from each other. In general, beta mapping-based VaR models that consider also the firm-specific risks tend to produce equal VaR measures with the traditional model. The plain beta VaR models that neglect the firm-specific risks appear to provide competent risk measures only for relatively large portfolios including at least 100 equities. Mapping onto the market seems to be preferable to the mapping onto the sectors.

Key words

Value at Risk (VaR), beta mapping, testing, backtesting, portfolio theory

Table of contents

1. Introduction 1

1.1. Motivation for study 1

1.2. Objectives and scope of study 2

1.3. Structure of thesis 3

2. VaR and variance-covariance approach 4

2.1. VaR concept 4

2.2. Past VaR accuracy findings 5

2.3. Variance-covariance approach 9

2.3.1. Problems in estimating covariance matrix 13

2.3.2. Basics of position mapping 14

3. Methodology 16

3.1. Beta mapping-based VaR models 16

3.1.1. Mapping onto market index and respective VaR models 16

3.1.2. Mapping onto sector indices and respective VaR models 20

3.2. Determination of equity and sector weights 22

3.3. Calculation of returns 23

3.4. Estimation of variances 25

3.4.1. Moving average (MA) method 26

3.4.2. Generalised autoregressive conditional heteroskedastic (GARCH) model 27

3.4.3. Exponentially weighted moving average (EWMA) method 28

3.4.4. Implied volatility approach 29

3.4.5. Comparison of methods and selection of appropriate method 32

3.5. Estimation of covariances 38

3.6. Estimation of betas 41

3.7. Tests and corresponding hypotheses 41

3.7.1. *t*-test for zero mean portfolio return 42

3.7.2. Anderson-Darling test for normality of portfolio returns 43

3.7.3. Kupiec tests for exceeded VaR measures 46

4. Data and portfolio characteristics 51

4.1. Data 51

4.2. Portfolio characteristics 52

4.2.1. Small portfolio 54

4.2.2. Medium portfolio 55

4.3.3. Large portfolio 57

5. Results 59
5.1. Standard deviation comparisons 59
5.1.1. Small portfolio 60
5.1.2. Medium portfolio 62
5.1.3. Large portfolio 64
5.2. VaR accuracy results 66
5.2.1. Traditional model 67
5.2.2. Beta model 71
5.2.3. Diagonal beta model 74
5.2.4. Sector-beta model 76
5.2.5. Diagonal sector-beta model 77
5.3. Summary results 80
6. Conclusions 85
References 88
Appendix A. Comparison of representative mapping approaches 91
Appendix B. Small portfolio's return figures 93
Appendix C. Medium portfolio's return figures 94
Appendix D. Large portfolio's return figures 95
Appendix E. Small portfolio's supplementary STD figures 96
Appendix F. Medium portfolio's supplementary STD figures 97
Appendix G. Large portfolio's supplementary STD figures 98

1. Introduction

1.1. Motivation for study

The Value at Risk (VaR) approach has been increasingly popular in the field of market risk management since the J.P. Morgan's public launch of its RiskMetrics™ system in 1994. VaR is nowadays widely used by financial institutions, nonfinancial corporations, and institutional investors to measure downside market risk of their portfolios. (Linsmeier and Pearson, 2000)

One common approach to calculate VaR measures for equity portfolios is based on portfolio theory. This approach, generally called variance-covariance approach, requires computation of the covariance matrix of the portfolio returns. Unfortunately, as the portfolios become relatively large, it's often necessary to make some simplifying assumptions concerning the covariance matrix; otherwise it might be impossible to make VaR calculations at all or the calculations might be considerably time-consuming. When large equity portfolios are concerned, one proposed approximation to the covariance matrix, which is also used by RiskMetrics, is based on mapping the individual equity positions onto the market index (RiskMetrics - Technical Document, 1996). Variance and covariance estimates of the covariance matrix are then derived using the beta coefficients and the estimate of the variance of the market index.

Supposedly the study of Johansson, Seiler and Tjarnberg (1999) is the only published study concerning the accuracy testing of beta mapping-based VaR models. It's found in their study that beta mapping works reasonably well for highly diversified portfolios when VaR measures are calculated at the 95% confidence level. On the other hand, they conclude that beta mapping doesn't work so well for small, relatively undiversified portfolios. However, their data sample might include illiquid shares, which could easily result in biased and underestimated betas. Since the betas are the main determinants of the beta mapping-based VaR measures, the presence of illiquid equities in a portfolio would probably lead to too low risk measures, and thus also to the biased results concerning the accuracy of beta mapping in VaR framework. Also, Johansson, Seiler and Tjarnberg apply relatively short backtest period of only 261 business days, which might bias their results even more.

As many market participants use in their equity portfolio risk management, either directly in their own systems or indirectly through RiskMetrics, VaR models that are based on beta mapping, and as the study of Johansson, Seiler and Tjarnberg (1999) is rather insufficient, it's of great importance to further test how accurately beta mapping-based VaR models can capture market risk of an equity portfolio. Moreover, there can be a large number of institutions in the marketplace using lower VaR measures than they believe, since beta mapping may underestimate true VaR due to the simplifications it implies.

1.2. Objectives and scope of study

The main objective of this study is to find out how well do the four different beta mapping-based VaR models work on equity portfolios. Two of these models are based on mapping to the market index, while the other two map the individual equity positions into the sector indices. Accuracy of these VaR models is tested by making statistical analysis with historical data from the U.S. equity markets, particularly from the S&P 500 index companies, from years 1991-2000; due to the parameter estimation requirements, the actual backtesting period is slightly shorter and covers years 1995-2000. The S&P 500 constituent equities are chosen for the study, since they should be actively traded and thus liquid, which is very essential in order to get reliable results (S&P 500 Index Methodology, 2001). The four beta mapping-based VaR models are tested on their own and their performance is also compared with the traditional variance-covariance VaR model, in which the covariance matrix is estimated in a standard way by using individual equities' return data. Therefore, as a by-product of this study, it is further investigated how accurate VaR measures does the traditional variance-covariance VaR model produce for equity portfolios.

Another objective of this study is to be as practical as possible. Portfolio managers are perhaps more interested in weekly than daily VaR measures, depending on the level of their trading activities, so both weekly and daily VaR measures are tested. Furthermore, the VaR measures are calculated and tested for the two most commonly used VaR confidence levels of 95% and 99%. These dimensions add directly value to the study of Johansson, Seiler and Tjarnberg (1999), which considers only daily 95% VaR measures.

An important factor concerning the power of beta mapping-based VaR models to produce accurate VaR measures is presumably diversification. Therefore, in order to perceive the effect of diversification on accuracy of the beta mapping-based VaR models, portfolios with different numbers of equities are tested; more precisely, portfolios including 30, 50 and 100 equities are considered in the study.

As the VaR models' inputs, and therefore also outputs, may be significantly sensitive to the time horizons from which they are estimated, this study uses three different estimation periods lengths of past 250, 500, and 1000 days. In accordance with the modern practices, the returns are calculated geometrically. Furthermore, the variances and covariances are estimated using equally weighted moving averages, and the betas are estimated through the standard ordinary least squares (OLS) regression procedures.

1.3. Structure of thesis

The thesis is structured as follows. Chapter 2 introduces the VaR concept, presents past VaR measure accuracy findings, and describes the variance-covariance approach with its characteristics. Chapter 3 explains thoroughly the methodology applied in the study. The data and the portfolio compositions are presented in Chapter 4. Each VaR model's VaR measure accuracy results are reported alongside with the standard deviation comparisons in Chapter 5. Finally, Chapter 6 concludes the study.

2. VaR and variance-covariance approach

2.1. VaR concept

The concept and use of VaR is relatively recent. Major financial firms first used VaR in the late 1980s to measure the risk of their trading portfolios (Linsmeier and Pearson, 2000). Since then, the use of VaR has exploded. The public launch of the RiskMetrics system by J.P. Morgan in 1994 provided a great impulse to the growth. In these days, many market participants to measure downside portfolio risk use VaR.

VaR is a single, summary, and statistical measure of possible portfolio losses. Subject to the simplifying assumptions used in its calculation, VaR effectively aggregates all of the risks in a portfolio into a single number, which is easy to understand and suits well for the risk reporting purposes. The term VaR refers to a particular amount of money, the maximum expected loss over a given horizon at a given confidence level. VaR can be defined in terms of absolute loss, or in terms of loss relative to the expected portfolio profit. The former is simply the maximum expected loss measured from the current portfolio value: For example, if the given time is one week, the given confidence level 99%, and the portfolio's absolute VaR $20 million, then it's estimated that the odds that the portfolio will decline in value by more than $20 million within the next week are 1 in a 100. The relative VaR measure is obtained by adding the expected portfolio profit to the absolute VaR measure.

No consensus has been reached on the best way to implement VaR analysis, and the choice of an appropriate method is dependent on many circumstances, especially on characteristics of an asset portfolio. The three basic methods presented in literature to calculate VaR measures are variance-covariance approach, historical simulation and Monte Carlo simulation. The study's five VaR models apply the variance-covariance approach, and this approach is described in a general way in Chapter 2.3.[1]

In the VaR research, development of new VaR calculation methods and modification of the existing methods has received much more attention than testing of the methods, which

[1] The basic principles of the historical and Monte Carlo simulation methods can be found for instance from Dowd (1998).

is at least as important; in the end, a highly sophisticated VaR measure estimator is completely useless, if it doesn't work well in practice. Some empirical findings do however exist on the accuracy abilities of different VaR models, and these findings are presented in the next chapter.

2.2. Past VaR accuracy findings

Mahoney (1995) studies the VaR measure accuracy of the variance-covariance approach and historical simulation methods for currency and equity portfolios. His currency portfolios are constructed of eight major currencies for which he has daily exchange rates from May 1975 to July 1995. On the other hand, the equity portfolios are constructed of six equity indices, which all represent separate currencies, and the respective dataset includes daily returns from October 1986 to July 1995. In the variance-covariance approach, Mahoney estimates the variances using equal and exponential weighting. He uses past 1000 days in the estimation of the equally weighted variance and in the historical simulation method. Thus, backtesting is applied for the both datasets only from the 1001st observations onwards. Mahoney concludes that for the currency as well as for the equity portfolios, the both VaR methods produce unbiased VaR measure estimates for the 90% and 95% confidence levels. However, only the historical simulation method seems to yield accurate VaR measures also for the higher confidence levels.

Hendricks (1996) does the similar study to Mahoney (1995), but only for currency portfolios. His dataset comprises also eight major currencies and their corresponding daily exchange rates for years 1978-1994. Due to the parameter estimation requirements, his actual backtest period for the VaR measures is a bit shorter and covers years 1983-1994. Hendricks uses equally and exponentially weighted moving average methods when estimating the portfolio variances in the variance-covariance approach. Also, different estimation and observation period lengths are applied in Hendricks' study. His findings are rather identical with Mahoney's findings. In other words, Hendricks finds that both the variance-covariance approach as well as the historical simulation methods produce in general accurate 95% VaR measures and the differences are rather insignificant among the different VaR and variance estimation methods. However, when the 99% VaR measures are concerned, especially the variance-covariance approach produces generally too low risk measures and fails to achieve the desired coverage level; Hendricks reports that accuracy of the historical simulation's 99% VaR measure seems to be dependent on the observation period length and only relatively long periods of about five years can produce valid VaR measures for the 99% confidence level.

Jackson, Maude and Perraudin (1997) complement the previously presented two VaR accuracy studies for the variance-covariance approach and historical simulation. Their test portfolios are composed jointly of currencies, interest rate products and equities, and thus seek to represent realistic asset portfolios actually held by market participants. They use approximately six-year long backtest period from June 1989 to April 1995. In addition to the daily returns, they consider also 10-day returns in their study. Consistently with the earlier studies, they find that the variance-covariance approach yields inaccurate and too low 99% daily VaR measures while historical simulation is capable to produce rather accurate measures when a relatively long estimation period length of past 24 months is applied. When the ten-day portfolio returns are considered, the variance-covariance approach produces however generally quite accurate 99% VaR measures, which are in general superior to the corresponding measures of the historical simulation method. Furthermore, the findings of Jackson, Maude and Perraudin indicate that the longer the estimation or observation period length the more accurate the VaR measure.

Polasek and Pojarliev (2000) investigate accuracy performance of GARCH models in the variance-covariance VaR approach. They use the NASDAQ 100 index as a proxy for the equity portfolio, and their backtesting data comprises of 396 daily return observations between 26th of September 1998 and 28th of April 2000. They find that GARCH models produce variance estimates that lead to accurate 95% VaR measures. Furthermore, they conclude that the GARCH variance estimates imply more accurate VaR measures than variances which are estimated applying either exponentially weighted moving average method or the equally weighted moving average method with past 800 observations.

Also Wong, Cheng and Wong (2001) examine how accurate VaR measures the variance-covariance approach yields for an equity portfolio when GARCH procedures are applied in the variance estimation. They use the Australia's All Ordinary Index (AOI) for their equity portfolio and their sample covers daily return data of the index from February 1983 to June

1999, implying a total of 4000 daily return observations for the backtesting. Slightly contrary to Polasek and Pojarliev (2000), Wong, Cheng and Wong conclude that application of GARCH-based variance estimation is not a reliable way to manage or measure the market risk of an equity portfolio, since the corresponding 99% VaR measures consistently underestimate the true market risk of the portfolio. However, this inconsistency between the findings may well be due to the limitations of the normal distribution to correctly estimate very extreme return outcomes. In other words, the GARCH variance estimates may in reality be in the both studies as accurate as possible, and the VAR measure shortcomings in the study of Wong, Cheng and Wong might result solely from the misspecified VaR method.

Sarma (2001) studies how accurate 95% and 99% VaR measures can be produced by applying the extreme value theory (EVT), which in contrary to the three common VaR calculation methods deals directly with the behaviour of the return distributions' tails. The key insight of EVT is that an unknown distribution of extreme returns asymptotically converges to a limiting distribution of a particular known form (Dowd, 1998). For the EVT-based VaR measures, Sarma applies the peaks-over-threshold (POT) model, which is presented in McNeil and Frey (2000).[2] In addition to the POT model, Sarma tests simultaneously the accuracy of the variance-covariance approach and historical simulation methods; variances are estimated applying GARCH procedures. He calculates and backtests 1312 daily VaR measures for a long and a short position in the Nifty equity portfolio from 8th of May 1996 to 13th of August 2001. Sarma concludes that for the both portfolio positions, all the three methods seem to produce accurate 95% and 99% VaR measures. He also ranks the three models based on their backtesting performance, but these rankings are rather dependent on the particular VaR confidence level and the position type. For instance, Sarma finds that the EVT-based POT model produces the most accurate 99% VaR measures for the long Nifty portfolio, while the variance-covariance approach is found to be best for the short Nifty portfolio, for the both 95% and 99% VaR measures.

---

[2] In brief, the POT model provides a framework for estimating the positive or negative tails of the return distribution by estimating what is known as the distribution of excesses over certain threshold point, which identifies the starting point of the tail. For specific information on the framework, see for example McNeil and Frey (2000).

Although VaR is used primarily for portfolios consisting of financial assets, no restrictions exist for its use for other kind of portfolios, too. As a consequence, Manfredo and Leuthold (2001) study whether the VaR is an effective tool in agricultural risk management. More precisely, they examine is the VaR analysis a competent way to measure market risk of the cattle feeding margin, which is defined as the difference between fed cattle prices (output price) and prices of corn and feeder cattle (input prices). Thus, the components of the cattle feeding margin, that is the fed cattle, feeder cattle, and corn prices, serve as a portfolio of assets. Manfredo and Leuthold consider weekly 90%, 95% and 99% VaR measures and their backtest period is from January 1987 to October 1997, providing a total of 564 weekly VaR measures and portfolio outcomes. They apply the variance-covariance approach and historical simulation methods for the VaR measures; variances are estimated in the variance-covariance approach by equally and exponentially weighting past observations, by using GARCH methods, and by using implied volatilities of option prices. They conclude that for each VaR confidence level, the both VaR calculation methods produce in general well-calibrated VaR measures. Different variance estimation methods have however slightly differing accuracy levels in the variance-covariance approach. The most accurate VaR measures are obtained by exponentially weighting past return outcomes, while the equal weighting is the second most accurate. On the other hand, application of the GARCH and implied variances result in a bit weaker VaR measures mainly due to the problems in estimating correlations and covariances consistently with the individual variances. Overall, the findings of Manfredo and Leuthold support the idea that the variance-covariance approach as well as historical simulation methods can produce accurate VaR measures also for portfolios, which consist of other than financial assets.

The findings presented above clearly evidence that accuracy of a VaR measure is not only sensitive to the selected VaR calculation method, but also to the chosen VaR confidence level and time horizon as well as on to the portfolio under consideration. Generally speaking, different VaR calculation methods seem to provide rather accurate VaR measures for an equity portfolio, especially at the 95% confidence level. Nevertheless, the previously reported past VaR measure accuracy findings give little nor any guidance whether or not the widely applied beta mapping-based VaR models lead to accurate downside market risk measures.

To clarify this economically important issue, Johansson, Seiler and Tjarnberg (1999) examine the accuracy of daily 95% VaR measures produced applying beta mapping. They apply beta mapping in the variance-covariance approach and in the Monte Carlo simulation methods. In addition to this, they also test the accuracy performance of the traditional variance-covariance approach as well as the historical simulation methods. Variances in the three parametric methods are estimated applying equal as well as exponential weighting of past return observations. In the variance estimation and in the historical simulation method, four different estimation and observation period lengths are applied; precisely, past 89, 261, 524, and 1048 days. From 62 randomly chosen equities from four different countries – Canada, Germany, Japan and the United States – they compose three progressively diversified portfolios with 7, 15 and 62 equities. In their study, they use five-year long daily return data from 10th of July 1993 to 10th of July 1998, but due to the estimation requirements the actual backtesting period is relatively short and covers only 261 days. Based on their findings concerning the study's most effectively diversified portfolio with its 62 equities, Johansson, Seiler and Tjarnberg conclude that application of beta mapping for highly diversified portfolios results in reasonable VaR measures, which are also competent with the VaR measures estimated applying the common methods of variance-covariance approach and historical simulation. However, for smaller, less diversified portfolios, beta mapping doesn't seem to be appropriate, while the two common methods maintain their ability to yield accurate daily 95% VaR measures. Furthermore, the findings of their study support that relatively short estimation and observation period lengths should be applied in VaR calculations, which is in contrast to the earlier findings of Jackson, Maude and Perraudin (1997). Finally, although the research of Johansson, Seiler and Tjarnberg gives valuable information concerning the accuracy of beta mapping in VaR calculations, it's nevertheless inadequate and leaves important questions unanswered, as pointed out in several contexts throughout this study.

2.3. Variance-covariance approach

The variance-covariance approach of calculating VaR measures is based on the estimate of the variance-covariance matrix of asset returns. Usually the estimation of the matrix is made by using historical time series of asset returns and calculating their variances and covariances, but also other methods exist for the estimation, such as estimating variances using implied volatitilities of option prices.

The main assumption of the variance-covariance approach is that the distribution of the portfolio return is normal. If the individual returns of the assets in the portfolio are normally distributed, then naturally also the portfolio return is normally distributed. But the normality assumption could hold even if the individual returns are not normally distributed, since the central limit theorem of statistics indicates that if the portfolio is fairly well diversified and individual returns are sufficiently independent of each other, the portfolio return converges to a normal distribution. Thus, the normality assumption concerning the portfolio return is often considered relatively reasonable. Especially in the case of equity portfolios, where all individual positions are expected to be linear in the underlying risk factors, the normality assumption is usually considered adequate.

Portfolio's relative VaR measure is calculated rather simply through Equation (1) below.[3] The minus sign at the beginning of the equation stems from the fact that VaR is usually expressed as a positive number.

$$VaR_p = -\alpha\sigma_p W \tag{1}$$

where

α = confidence level parameter (e.g. -1,65 for 95% and –2,33 for 99% confidence level),

σ_p = standard deviation of portfolio return,

W = initial portfolio value.

If one would instead like to calculate portfolio's absolute VaR measure, the calculation would be carried out using the following equation

$$VaR_p = -\alpha\sigma_p W - \mu_p W \tag{2}$$

where

μ_p = portfolio's expected return.

[3] For a detailed derivation of the equation, see for example Jorion (2000).

Comparing equations (1) and (2) it can be observed that the relative VaR measure is easier to calculate because it doesn't require that the expected portfolio return μ_p is known. This is why it's more convenient to work with the relative VaR measures, and in all what follows in this study the term VaR refers to the relative VaR, if not otherwise noticed.

As can be seen from (1), the standard deviation of the portfolio return, that is the square root of the portfolio return variance, is the only parameter that must be estimated in order to calculate the portfolio's VaR measure. Thus, the standard deviation should reflect the selected VaR time horizon: For instance for the daily VaR measure, the daily standard deviation is needed. However, it's not necessary to estimate totally separately a standard deviation for all different VaR time horizons if some simplifying assumptions are made concerning the asset returns, as will be shown later in Chapter 3.4.5.

Portfolio theory solves the problem of calculating the portfolio standard deviation. One of the core assumptions underlying in the portfolio theory's solution is that the portfolio return r_p is the sum of the weighted returns of the individual assets r_i. The weights are constructed to sum to unity by scaling the monetary positions in each asset W_i by the portfolio total market value W. Mathematically, the portfolio return is calculated in portfolio theory as shown below. (Sharpe, 2000)

$$r_p = \sum_{i=1}^{N} w_i r_i \tag{3}$$

where

N = number of assets in portfolio,

w_i = weight of *i*th asset in portfolio.

Equation (3) holds for the arithmetic returns, but it's only an approximation for the geometric returns. This is due to the fact that the logarithm of a sum is not the same as the sum of logarithms. Based on this, it may seem to be feasible to use arithmetic calculation of returns. However, according to Campbell, Lo and MacKinlay (1997) the approximation error of the geometric returns in Equation (3) is usually negligible concerning empirical applications. Especially in daily and weekly returns, which are generally rather small, the approximation error of the geometric return calculation is expected to be very marginal. Moreover, the use of the geometric returns offer some meaningful advantages over the arithmetic returns.[4] Therefore, it's a common practice to use geometric returns in research, and this study doesn't depart from this custom.

The portfolio variance depends not only on the variances of the individual series but also on the covariances between two series. As two different series aren't generally perfectly correlated and their covariance isn't simply the product of their respective variances, the diversification of the portfolio risk is beneficial; the smaller the correlation, the greater the benefit. Portfolio theory introduces the following formula for the calculation of the portfolio variance (Sharpe, 2000)

$$\sigma_p^2 = \sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i=1}^{N} \sum_{j=1, j \neq i}^{N} w_i w_j \sigma_{ij} \tag{4}$$

where

σ_i^2 = variance of return of *i*th asset,

σ_{ij} = covariance between returns of *i*th and *j*th assets.

As the number of assets increases, it becomes difficult to keep track of all covariance terms, which is why it's more convenient to use matrix notation. The variance of the portfolio return can be written as

$$\sigma_p^2 = \left[w_1 \cdots w_N \right] \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1N} \\ \vdots & \ddots & \vdots \\ \sigma_{N1} & \cdots & \sigma_{NN} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix}. \tag{5}$$

Defining w as the weight vector and Σ as the covariance matrix of the returns, the variance of the portfolio return can be written compactly as

$$\sigma_p^2 = w' \Sigma w . \tag{6}$$

[4] For discussion on differences between arithmetic and geometric returns, see Chapter 3.3.

VaR Equation (1) can now be rewritten in a relatively compact form, which includes also the calculation of the portfolio standard deviation. The arising VaR model is called Traditional model in this study and it's of the following form

$$VaR_p = -\alpha\sqrt{w'\Sigma w}W \, . \quad (7)$$

Thus, as noted earlier, in order to calculate the portfolio VaR applying the variance-covariance approach, the covariance matrix must be estimated. Although it may at first seem to be an easy task to calculate the portfolio VaR using Equation (7) above, the implementation is not always so straightforward. Especially when the portfolios become large, the estimation of the covariance matrix usually raises problems, which are discussed in the next subchapter.

2.3.1. Problems in estimating covariance matrix

Large amount of data concerning the variance and covariance terms is needed for the calculation of the covariance matrix. If the portfolio includes N different assets, data would be needed on N separate variance terms, one for each asset, plus also data on $N(N-1)/2$ covariance terms, a total altogether of $N(N+1)/2$ pieces of information. As new assets are added to the portfolio, the additional amount of covariance data needed grows geometrically. Clearly, as N gets large, the amount of data needed becomes enormous. In practice, it easily becomes impossible to collect and process such data for any but a small proportion of possible assets. (Dowd, 1998)

It's normally reluctant to work with very high dimension covariance matrices, even if all the data were readily available. To guarantee that the estimated variance of the portfolio return, which is calculated through Equation (6), is non-negative, it's required that the corresponding covariance matrix Σ is positive semi-definite. If Σ is positive semi-definite, the portfolio variance and the respective standard deviation are always positive with any nonzero weight-vector w in Equation (6) (Simon and Blume, 1994).

According to Dowd (1998) the condition of positive semi-definite matrix is satisfied if two other conditions hold, both of which put limits on the size of the covariance matrix.

Firstly, the number of observations from which the covariance matrix is estimated must be at least as large as the number of dimensions on the matrix itself. For example, if the variance and covariance terms are calculated with 100 observations, the first condition requires that the covariance matrix has a dimensionality of no more than 100. Secondly, none of the included time series can be linearly correlated with other series or group of series. In other words, each series must have some independent movement of its own, distinct from the movements of other series. It's usually the latter condition that mostly constrains the covariance matrix calculations. In practical terms, problems frequently arise when the portfolio includes groups of assets that, whilst not perfectly correlated, are sufficiently closely correlated with each other, so that the estimated covariance matrix fails to be positive semi-definite due to the rounding errors in correlations.

A method is needed to overcome the lack of positive semi-definiteness associated with large covariance matrices. The usual solution is to scale down the dimensionality of the matrix by mapping the portfolio's assets onto benchmark assets for which the required data is available. The next section introduces the basics of how mapping can actually be done.

2.3.2. Basics of position mapping

By mapping positions, it's possible to estimate VaR measures, although with some approximation error, in situations that would otherwise be informationally extremely demanding or even impossible. Mapping does not only restore positive semi-definity of the covariance matrix by scaling down the dimensionality, but it also reduces the amount of noise in calculations and speeds up computations (Dowd, 1998). Jorion (1996) argues that these are considerable benefits, and that's why the design of the VaR system, including the number of variables that need to be estimated, is central. Mapping can be done in two ways, in a quantitative and in a representative way.

In the quantitative mapping, the key factors of the covariance matrix are identified by principal components analysis (PCA) or factor analysis (FA). These both procedures seek to identify the independent sources of movement within a group of time series, such as historical return data. Usually, only a relatively small number of principal components or factors need to be constructed to explain most of the movement. Once the principal components or factors are known, each individual asset in the portfolio will be mapped on these. Each principal component or factor is constructed to be independent of the others, so all have zero covariance with each other and the only non-zero elements of the principal components or factor covariance matrix will be the variances. Thus, principal components and factor analysis can drastically cut down the number of parameters, which need to be estimated for the covariance matrix. For example, large proportion of bond price movements can generally be approximated using three principal components, and then only three parameters, variances of these principal components, would be needed for the covariance matrix. (Dowd, 1998)

In the representative mapping, a set of core assets that can be regarded as representative of the assets actually held is selected. The objective is to have a rich enough set of core assets to be able to achieve efficient proxies for the assets in the portfolio whilst not having so many core assets that the problem of high dimensionality is faced again. After the selection of the core assets, the data on their variances and covariances is collected. Then synthetic substitute for each asset in the portfolio is derived by mapping each asset onto those core assets, which most effectively represent the asset; for example, in the case of equities, a common way is to map each individual equity position onto equity index using the beta coefficients. Finally, the portfolio VaR is calculated by using the synthetic substitutes. Exhibit 1 below summarises the central features of position mapping in VaR environment. The results of this study will eventually indicate whether or not the advantages outweigh the disadvantages.

Exhibit 1.
Features of mapping in VaR calculations

Advantages	Disadvantages
• Speeds up calculations	• May miss relevant information
• Ensures nonnegative VaR measures	• Examination of most appropriate principal components, factors or core assets can be problematic
• Lowers data requirements	
• Can make otherwise impossible VaR calculaions possible	

3. Methodology

In order to ensure that the study is as transparent as possible, the methodology applied in achieving the study's VaR measure accuracy results is presented in detail alongside justifications. Firstly, the study's beta mapping-based VaR models and their foundations are introduced. Since the output of any VaR model doesn't depend only on its design but also on the inputs, the subsequent subchapters present how weights, returns, variances, covariances and betas are defined and calculated in the study. Each of the study's five VaR models could yield very differing risk measures, when different kinds of methods are applied especially in the variance and covariance estimation. Because of this high dependence or sensitivity, all the common variance and covariance estimation techniques as well as their characteristics are presented. Also, their suitability for this particular study is discussed so that the most appropriate variance and covariance estimation methods can be selected. Finally, the last subchapter presents the tests applied in the study and the corresponding hypotheses.

3.1. Beta mapping-based VaR models

The study's four VaR models that apply mapping procedures are all based on representative mapping. Two of the four models base on mapping the equity positions onto market index while the other two are based on mapping onto sector indices. These both mapping approaches and the respective VaR models are explained in detail in the following two subchapters.

3.1.1. Mapping onto market index and respective VaR models

The assumption in mapping equity positions onto the market index is that the common movement in all equity positions is due to one common factor only, the market. Particularly, it's assumed that individual equity returns can be estimated effectively using the market model. The model is completely statistical in contrast to economic models, such as the Capital Asset Pricing Model (CAPM), which restrict the parameters of statistical models. The market model relates the return of an equity to the return of the market portfolio.[5] The model's linear specification follows from the assumed joint normality of asset returns. The model is as follows

$$R_i = \alpha_i + \beta_i R_m + \varepsilon_i \tag{8}$$

where

R_i = return on equity *i*,

α_i = firm-specific constant,

β_i = beta coefficient of equity *i*,

R_m = return on market index,

ε_i = residual of equity *i*.

In addition, the following assumptions concerning the residual are made

$$E[\varepsilon_i] = 0 \qquad E[\varepsilon_i R_m] = 0 \qquad E[\varepsilon_i \varepsilon_j] = 0 \qquad E[\varepsilon_i^2] = \sigma_{\varepsilon,i}^2. \tag{9}$$

Thus, it's assumed that the residual ε_i is not correlated with the market or across equities. Variance of the return on equity can then be derived as follows

$$\sigma_i^2 = \mathrm{var}(\beta_i R_m + \varepsilon_i) = \beta_i^2 \ \mathrm{var}(R_m) + \mathrm{var}(\varepsilon_i) + 2\,\mathrm{cov}(\beta_i R_m, \varepsilon_i) = \beta_i^2 \sigma_m^2 + \sigma_{\varepsilon,i}^2. \tag{10}$$

From (10) it can be seen that the variance of an individual equity position consists of a market-based component $\beta_i^2 \sigma_m^2$ and a firm-specific component $\sigma_{\varepsilon,i}^2$. The covariance between two equities is

$$\sigma_{ij} = \mathrm{cov}(\alpha_i + \beta_i R_m + \varepsilon_i, \alpha_j + \beta_j R_m + \varepsilon_j) = \mathrm{cov}(\beta_i R_m, \beta_j R_m) = \beta_i \beta_j \sigma_m^2. \tag{11}$$

[5] The specification actually requires that the composition of the market portfolio and the equity weights in the portfolio remain unchanged. In applications a broad-based equity index, such as the S&P 500 index, is generally used for the market portfolio. Thus, in reality the compositions and weights evolve over time. However, the changes over time are small enough that they have little effect on empirical work. (Campbell, Lo and MacKinlay, 1997)

When individual equity positions are mapped onto the market index, Equation (11) effectively implies that the covariance between two equities is solely due to the common factor, market. Thus, the full covariance matrix is

$$\Sigma = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \end{bmatrix} \begin{bmatrix} \beta_1 & \cdots & \beta_N \end{bmatrix} \sigma_m^2 + \begin{bmatrix} \sigma_{\varepsilon,1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{\varepsilon,N}^2 \end{bmatrix}. \tag{12}$$

Written compactly in matrix notation, the covariance matrix is

$$\Sigma = \beta\beta'\sigma_m^2 + D_\varepsilon . \tag{13}$$

Applying mapping onto the market index, only a total of $2N + 1$ parameters need to be estimated for the covariance matrix; in other words, N parameters for betas, N for diagonal matrix D_ε, and one for σ_m^2. Compared to the standard situation, where $N(N + 1)/2$ terms need to be estimated for the covariance matrix, the reduction is significant. For instance, with 100 assets the number of parameter estimates is reduced from 5050 to 201, which is a meaningful improvement.

Mapping individual equity positions onto the market index simplifies the variance of the portfolio return to the following form

$$\sigma_p^2 = w'\Sigma w = (w'\beta\beta'w)\sigma_m^2 + w'D_\varepsilon w = (\beta_p\sigma_m)^2 + w'D_\varepsilon w . \tag{14}$$

Furthermore, the variance of large, well-diversified portfolio simplifies even further, reflecting only exposure to the common factor, the market. The latter term $w'D_\varepsilon w$ in Equation (14) consists of $\sum_{i=1}^{N} w_i^2\sigma_{\varepsilon,i}^2$, which becomes very small as the number of equities in the portfolio increases. For instance, if all the residual variances are assumed to be identical and each equity has an equal weight in the portfolio, this latter term is

$\left[\sum_{i=1}^{N}(1/N)^2\right]\sigma_\varepsilon^2 = (1/N)\sigma_\varepsilon^2$, which converges to 0 as N increases.[6] Therefore, the variance of the portfolio converges to

$$\sigma_p^2 \rightarrow (w'\beta\beta'w)\sigma_m^2 = (\beta_p\sigma_m)^2. \tag{15}$$

In Formula (15) above, the portfolio variance is only due to the common factor, the market. Thus, in large portfolios, firm-specific risk becomes unimportant for the purpose of measuring VaR, and the number of parameters required reduces to $N+1$.

The VaR calculations can be made somewhat more rapidly and easily, if Formula (15) is used in estimating the portfolio variance instead of Formula (14). Johansson, Seiler and Tjarnberg (1999) test only beta mapping-based VaR models that apply Formula (15) in the portfolio variance estimation. However, some relevant information may get lost and the variance and therefore also the VaR may be underestimated, if the portfolio is not well enough diversified and Formula (15) is applied. Due to this, VaR models based on the both variance calculation methods, (14) and (15), are tested in the study with all the three portfolios. Portfolios including different numbers of equities are tested in order to get indication how the relevance of the diagonal matrix D_ε decreases as the portfolios get larger.

The study's two VaR models that apply mapping of individual equity positions onto the market index are shown below; these are Beta model (16) and Diagonal beta model (17). Beta model, which considers only the general market risk, is exactly the same model as RiskMetrics uses for equities (RiskMetrics - Technical Document, 1996). Diagonal beta model considers also the firm-specific risk.

$$VaR_p = -\alpha\beta_p\sigma_m W \tag{16}$$

$$VaR_p = -\alpha\sqrt{(\beta_p\sigma_m)^2 + w'D_\varepsilon w}W \tag{17}$$

[6] Under these conditions, Sharpe (2000) states that a portfolio containing at least 15 assets can be considered well diversified; in other words, most of the non-systematic, firm-specific risk is diversified away.

3.1.2. Mapping onto sector indices and respective VaR models

Instead of using the general market index as the common factor driving the returns of individual equities, sector indices can also be used. A sector index is expected to explain the common movement of equities in the particular sector at least as well as the general market index. The model (8) is modified to the following form

$$R_i = \alpha_i + \beta_i R_{Si} + \varepsilon_i \tag{18}$$

where

R_{Si} = return on equity *i*'s respective sector index.

The residual assumptions presented in (7) are expected to apply also for the model (18). When mapping equity positions onto the sector indices, the variance of individual equity gets very similar to the case of mapping onto the market index and is derived as follows

$$\sigma_i^2 = \mathrm{var}(\beta_i R_{Si} + \varepsilon_i) = \beta_i^2 \,\mathrm{var}(R_{Si}) + \mathrm{var}(\varepsilon_i) + 2\mathrm{cov}(\beta_i R_{Si}, \varepsilon_i) = \beta_i^2 \sigma_{Si}^2 + \sigma_{\varepsilon,i}^2 . \tag{19}$$

On the other hand, the covariance between two equities is now

$$\begin{aligned} \sigma_{ij} &= \mathrm{cov}(\alpha_i + \beta_i R_{Si} + \varepsilon_i, \alpha_j + \beta_j R_{Sj} + \varepsilon_j) \\ &= \mathrm{cov}(\beta_i R_{Si}, \beta_j R_{Sj}) = \beta_i \beta_j \sigma_{SiSj} . \end{aligned} \tag{20}$$

If the both equities represent the same sector, the covariance in Equation (20) simplifies to $\beta_i \beta_j \sigma_{Si}^2$, a sector *i* variance multiplied by two different betas. Thus, the covariance terms are needed only for the equities, which represent different sectors. This means that equities can be grouped by sectors.

The equity grouping is made in the following manner: Firstly, each of the portfolio's equities is classified into one sector and its beta for the corresponding sector index is calculated. Then, for each sector a portfolio sector-beta is calculated by using the betas of those equities, which represent the particular sector, and their corresponding weights on that sector; in other words, a portfolio sector-beta is a weighted average of the corresponding sector's individual betas. Finally, the portfolio sector-betas can be used in calculating the portfolio variance. They can be understood as kind of correction terms, which correct the amount of risk that a specific portfolio has on different sectors. Using matrices, the variance of the portfolio can be written as follows

$$\begin{aligned}\sigma_p^2 &= \begin{bmatrix} w_{S1}\beta_{S1} & \cdots & w_{Sn}\beta_{Sn} \end{bmatrix} \begin{bmatrix} \sigma_{S1S1} & \cdots & \sigma_{S1Sn} \\ \vdots & \ddots & \vdots \\ \sigma_{SnS1} & \cdots & \sigma_{SnSn} \end{bmatrix} \begin{bmatrix} w_{S1}\beta_{S1} \\ \vdots \\ w_{Sn}\beta_{Sn} \end{bmatrix} \\ &+ \begin{bmatrix} w_{S1} & \cdots & w_{Sn} \end{bmatrix} \begin{bmatrix} \sigma_{\varepsilon,S1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{\varepsilon,Sn}^2 \end{bmatrix} \begin{bmatrix} w_{S1} \\ \vdots \\ w_{Sn} \end{bmatrix}\end{aligned} \tag{21}$$

where

n = number of different sectors represented in portfolio,

w_{Si} = weight of sector i in portfolio,

β_{Si} = portfolio beta of sector i,

σ_{SiSj} = covariance between returns of ith and jth sector indices,

$\sigma_{\varepsilon,Si}^2$ = firm-specific variance of sector i.

The firm-specific variance of sector i $\sigma_{\varepsilon,Si}^2$ is composed of $\sum_{i=1}^{m} w_i^2 \sigma_{\varepsilon,i}^2$, where the weights and the firm-specific variances represent the portfolio's equities on the particular sector.

The portfolio variance presented in (21) can be defined more compactly using the matrix notation as follows

$$\sigma_p^2 = w'_{S\beta} \Sigma_S w_{S\beta} + w'_S D_{\varepsilon S} w_S. \tag{22}$$

Similarly to the previous chapter's approach that maps equity positions onto the market index, two VaR models are considered in the study that apply position mapping onto the sector indices. These two models, Sector-beta model (23) and Diagonal sector-beta model

(24), are shown below. Sector-beta model takes into account only the general risks of different sectors, while Diagonal sector-beta model considers also the firm-specific risks.

$$VaR_p = -\alpha\sqrt{w'_{S\beta}\Sigma_S w_{S\beta}}W \tag{23}$$

$$VaR_p = -\alpha\sqrt{w'_{S\beta}\Sigma_S w_{S\beta} + w'_S D_{\varepsilon S} w_S}W \tag{24}$$

The number of parameters, which need to be estimated, is slightly larger when mapping is applied onto the sector indices instead onto the market index. For a Diagonal sector-beta model, a total of $2N + n(n + 1)/2$ parameters are needed; that is N for betas, N for diagonal matrix $D_{\varepsilon S}$, and $n(n + 1)/2$ for the covariance matrix of the sector indices Σ_S. In the case of Sector-beta model, a total of $N + n(n + 1)/2$ terms need to be estimated. Thus, for instance for the study's largest portfolio that includes altogether 100 equities on eight different sectors, 236 parameters have to be estimated if mapping is applied onto the sector indices and also the firm-specific risks are considered in VaR calculations, while the corresponding number of required parameters is only 201 when the representative mapping is made onto the market index.

Although the estimation burden is bigger when the positions are mapped onto the sector indices instead onto the market index, the possible advancements in accuracy may outweigh this. As shown in Appendix A, the correlation conditions will effectively determine which of the two mapping approaches produces higher and thus presumably more accurate portfolio variances for VaR calculations. Shortly stated, if the different sector indices don't correlate enough with each other, the two VaR models applying mapping onto sector indices can imply lower portfolio variances and market risk measures than the corresponding two VaR models that apply mapping onto the market, even though individual equities would correlate more with sector indices than with the market index.

3.2. Determination of equity and sector weights

In this study the composition of each portfolio is fixed during the test period. In other words, it's assumed that no trading occurs for any of the three portfolios during six years, which is clearly a very unrealistic assumption. However, simulation of the portfolio changes through time in a realistic manner is a rather infeasible task, so the assumption concerning the fixed portfolio compositions has to be made for practical reasons.

Even though the portfolio compositions are fixed, the weights of the underlying equities and sectors do evolve according to the equity returns. A weight of an equity during the test period 1995-2000 is calculated by dividing its prevailing value in a portfolio by the prevailing total portfolio value. A sector weight in the portfolio is simply the sum of those equity weights, which represent the particular sector.

The initial amounts invested in portfolios' different equities, and their corresponding weights in the portfolios, are determined with the help of the equities' market capitalisations at the beginning of 1995. In determining the initial amounts and weights, the modified market capitalisations are used in order to avoid too large weight on a single equity. Because the equity price development during the test period is known for each equity in the portfolio, the market capitalisations at the beginning of 1995 are modified so that 10% is the maximum weight any single equity can have any time during the test period 1995-2000.

If no modification is made to the market capitalisations at the beginning of the test period, some equities have weights well over 20% during the test period, which is against the common views on effectively diversified portfolios and thus inappropriate. For example, Intel's weight in the large portfolio is approximately 22.5% in the early 2000, if no modification is made to the initial market capitalisations. The maximum limit of 10% for an individual equity weight is believed to guarantee that diversification is on an appropriate level in each of the three portfolios, given the number of equities included in the particular portfolio.

3.3. Calculation of returns

Indices' and equities' rates of returns, or simply returns, can be calculated by two alternative methods. These common alternatives are the arithmetic return (25) and the geometric return (26), which are described in their mathematical forms below for a daily return on day t.

$$r_t = \frac{P_t + D_t - P_{t-1}}{P_{t-1}} \tag{25}$$

$$R_t = \ln \frac{P_t + D_t}{P_{t-1}} \tag{26}$$

where

P_t = closing value of index/price of equity on day t,

D_t = dividend payment for equity associated with ex-date t,

P_{t-1} = closing value of index/price of equity on previous business day.

Both of the return calculation methods have their own advantages and disadvantages. The biggest drawback of the geometric return calculation is its inability to calculate the portfolio return correctly as a linear combination of the returns on the underlying equities. In other words, as pointed out earlier in Chapter 2.3, the application of Formula (3) produces only an approximation for the portfolio return, when geometric returns are applied. However, the approximation error in daily and weekly returns that both are generally relatively low is expected to be insignificant.

In research, it's an established custom to use geometric return calculation. This is due to important advantages it offers over the arithmetic return calculation. Firstly, the arithmetic return has a very serious problem in that it's not symmetric. For example, if the equity price rises from \$100 to \$200, the arithmetic return would be 100%, but if it falls back from \$200 to \$100, the arithmetic return is not –100% but only –50%. As a result, the arithmetic return on the negative side cannot be below –100%, while on the positive side there is no limit on the return. The statistical implication of this is that returns are skewed in the positive direction and application of the normal distribution becomes inappropriate.

The second important advantage of using geometric return calculation is that it makes calculations for multiple periods much easier compared to the arithmetic return calculation. This advantage arises from the fact that the geometric returns can be summed. For example, the two-day geometric return is simply the sum of the two successive periodic geometric returns. This is proven in (27) below, where the dividends are for simplicity assumed to be zero.

$$R_{t,2} = \ln \frac{P_t}{P_{t-2}} = \ln P_t - \ln P_{t-1} + \ln P_{t-1} - \ln P_{t-2} = \ln \frac{P_t}{P_{t-1}} + \ln \frac{P_{t-1}}{P_{t-2}} = R_t + R_{t-1} \quad (27)$$

The corresponding two-period arithmetic return is more involved. As (27) shows, no rebalancing of positions takes place when the two-period geometric return is calculated. This is unfortunately not the case with the arithmetic returns, which correspond to the situation of a fixed investment, that is, where gains are withdrawn and losses are added back. (Jorion, 2000)

The previously mentioned two important advantages of the geometric returns are the grounds for calculating all returns in this study geometrically, by using Formula (26). Especially the latter advantage concerning the ability to calculate returns for multiple days by summing daily returns is very useful in the estimation of the weekly portfolio standard deviation, as described in the next chapter. Calculation of the weekly returns applying addition implicitly means that possible dividends are reinvested into the particular equity.

3.4. Estimation of variances

Each of the five VaR models that are tested in the study requires estimation of return variances. Estimates of firm-specific variances for Diagonal beta and Diagonal sector-beta models are found through Formulas (10) and (19), respectively, when estimates of betas, index variances and equity variances are known.

Four common methods exist for the variance estimation. Three of them, that is moving average (MA), generalised autoregressive conditional heteroskedastic (GARCH) model and exponentially weighted moving average (EWMA), are based on the historical return series data. The fourth estimation method is forward-looking and derives an estimate for the variance, or rather its square volatility, by using option prices. These alternative estimation methods are presented in the next four subchapters. Thereafter the methods are compared in the fifth subchapter that also includes selection of an appropriate method for this study.

3.4.1. Moving average (MA) method

As the name already implies, MA uses a moving window of fixed length for estimating variance. Thus, applying the MA approach, the variance is estimated in the following way

$$\sigma_t^{\,2} = \frac{\sum_{i=1}^{M} (R_{t-i} - \bar{R})^2}{M-1} \tag{28}$$

where

$\bar{R}$ = mean return in sample,

M = length of estimation period.

MA is the simplest one of the four alternative variance estimation methods. In the MA approach, every squared deviation from the mean has the same weight $1/M$, so it ignores the dynamic ordering of observations. The oldest observations in the estimation period may no longer be so relevant, but they still receive the same weights as the most recent ones. In other words, the MA approach assumes that the underlying true variance is constant, and so cannot accommodate any changes in the variance over time. In particular, it fails to allow for the variance clustering, which is a well-established phenomenon in the return series. (Dowd, 1998)

A second disadvantage of using the MA approach is the phenomenon called ghosting. For example, if there has been a relatively large return M days ago, dropping this return as the estimation period moves one day forward will substantially affect the variance estimate. The size of this effect depends obviously on the length of the estimation period. Anyway, the substantial change in the variance estimate is totally an artifact of the estimation period length. This feature is generally called ghosting, since there is no apparent reason for the change in the variance estimate. The power of the ghosting feature can be lessened using longer estimation periods. This may however raise new problems because longer estimation periods could more easily miss the underlying variation in true variance. Therefore, the MA approach leaves wholly unanswered the optimal choice for the estimation period length. (Jorion, 2000)

3.4.2. Generalised autoregressive conditional heteroskedastic (GARCH) model

A GARCH model tries to overcome the problems of the MA approach. It puts more weight on the most recent observations, and in that way allow for time-varying variance and variance clustering. The model assumes that the return variance follows a predictable process (Jorion, 2000). The generic GARCH model, Formula (29) below, presents a conditional variance estimator that depends on both p lagged values of squared returns and q lagged variance estimates. All parameters a_0, a_i:s and b_i:s are assumed to be positive. (Dowd, 1998)

$$\sigma_t^{\ 2} = a_0 + \sum_{i=1}^{p} a_i R_{t-i}^2 + \sum_{i=1}^{q} b_i \sigma_{t-i}^2 \tag{29}$$

The most common GARCH model is the GARCH(1,1) process, which depends only on the latest innovation, or return, and the previous conditional variance. Thus, mathematically defined the GARCH(1,1) variance estimator is simply

$$\sigma_t^{\ 2} = a_0 + a_1 R_{t-1}^2 + b\sigma_{t-1}^2 . \tag{30}$$

The average, unconditional variance of the GARCH(1,1) process can be found by setting $E(R_{t-1}^2) = \sigma_t^2 = \sigma_{t-1}^2 = \sigma^2$, and solving Formula (30) for σ^2. The following result is obtained

$$\sigma^2 = \frac{a_0}{1 - a_1 - b} . \tag{31}$$

For the GARCH(1,1) process to be stationary and allow for mean reversion that is usually considered reasonable variance behaviour, the parameter sum $a_1 + b$ must be less than one. This sum is also called the persistence, which determines how long a large shock will affect the conditional variance. High persistence means that the shock will decay slowly and the conditional variance will be high for a relatively long time. (Jorion, 2000)

Although GARCH models seem to be rather effective variance estimators in the sense that they can capture the observed variance clustering, their application isn't free from problems. The biggest drawback of these models is their nonlinearity. According to Jorion (2000) the parameters of a GARCH model have to be estimated by maximisation of the likelihood function, which involves a numerical optimisation.[7] In addition, Dowd (1998) argues that the GARCH models can be unstable, and therefore produce unreliable variance forecasts when forecasting out of the period, which is used to estimate the model's parameters.

3.4.3. Exponentially weighted moving average (EWMA) method

The EWMA approach can be viewed as a special case of the GARCH(1,1) process, in which the intercept a_0 is zero and the sum of the two remaining parameters is one. Thus, as is the case for the GARCH(1,1) process, in the EWMA model the whole history is summarised by only one number, σ_{t-1}^2. This is in contrast to the MA approach, for instance, where the last M returns must be used to construct the variance estimate. Formally, the EWMA variance estimate for time t is a weighted average of the previous variance estimate, using weight λ, and of the latest squared innovation, using weight 1 - λ

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)R_{t-1}^2. \tag{32}$$

The parameter λ is generally called the decay factor and must be less than one. The higher the decay factor the slower the weights of the past observations decay. Therefore, the number of the effective observations in the EWMA approach grows with the decay factor.

---

[7] Typically, researchers assume that the scaled residuals $\varepsilon_t = R_t / \sqrt{\sigma_t^2}$ have a normal distribution and are independent. If there are T observations, their joint density is the product of the densities for each time period t. The optimisation seeks to maximise the natural logarithm of the likelihood function $\max F(a_0, a_1, b \mid R) = \sum_{t=1}^{T} \ln f(R_t \mid \sigma_t^2) = \sum_{t=1}^{T}\left(\ln \frac{1}{\sqrt{2\pi\sigma_t^2}} - \frac{R_t^2}{2\sigma_t^2}\right)$, where f is the normal density function.

By recursively replacing σ_{t-1}^2 in Formula (32), the variance estimate can be written in the form of the exponentially weighted moving average that gives the EWMA model its name

$$\sigma_t^2 = (1-\lambda)\sum_{i=1}^{\infty} \lambda^{i-1} R_{t-i}^2 . \tag{33}$$

The EMWA approach is particularly easy to implement because it relies only on one parameter. Thus, it's more robust to estimation error than the generic GARCH models. Theoretically, the decay factor λ could be found from maximising the likelihood function. Operationally, this would however be a daunting task to perform every day for many return series. An optimisation has also other disadvantages. The optimal decay factor may vary not only across equities but also over time, thus losing consistency over different periods. In addition, different values of λ create contradictions for the EWMA covariance estimates and may lead to correlation coefficients greater than one.[8] (Jorion, 2000)

RiskMetrics estimates variances by applying the EWMA method. They find that the optimal decay factor for daily variance estimates is 0.94 and that the decay factor of 0.97 is optimal for monthly variance estimates. These decay factors are used continuously in the RiskMetrics' system for all equities and other assets. Thus, the daily and monthly variance estimators are inconsistent with each other.[9] However, according to Jorion (2000) the both models approximate the behaviour of the actual data quite well and are robust to misspecification. (RiskMetrics - Technical Document, 1996)

3.4.4. Implied volatility approach

The MA, GARCH and EWMA approaches have been criticised because they rely solely on the historical data. For instance, situations involving changes in market fundamentals are simply not reflected in recent historical data. In order to get estimates of the future variance or its square volatility that effectively take into account the future expectations in the market, the implied volatility approach can be applied.

---

[8] The EWMA covariance estimation method is presented on page 39.

[9] RiskMetrics uses Formula (32) introduced on page 28 for daily variance estimates and Formula (38) introduced on page 36 for monthly estimates.

The implied volatility approach is based on recovering the estimate of the future volatility from the observed option prices. Options are derivative instruments whose prices are influenced by a number of factors, all of which are directly observable from the market except the volatility of the underlying instrument's return. Therefore, the market's view on future volatility of the underlying instrument's return effectively determines the market value of the particular option. Thus, given the market price of an option, one can seek the implied return volatility of the underlying instrument by finding the volatility that matches the option pricing model's price with the market price.

Obviously, one should use an option pricing model that is widely used in the marketplace in order to avoid getting biased implied volatilities. For the European call options, for instance, the model originally introduced by Fischer Black and Myron Scholes in 1973, that is the BS model, is a standard valuation tool in the option markets. Thus, in the case of the European equity call option where dividends are assumed to be zero, the implied volatility is the yearly return volatility of the underlying equity σ, which satisfies the following equation for the call option's market price $C_{t,market}$, given current market data and the option's characteristics[10]

$$C_{t,market} = P_t N(d_1) - Ke^{-r_f(T-t)} N(d_2) \tag{34}$$

where

$$d_1 \quad = \frac{\ln\left(\frac{P_t}{K}\right) + \left(r_f + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}},$$

$d_2 \quad = d_1 - \sigma\sqrt{T-t}$,

$N(d_i)$ = cumulative standard normal distribution function evaluated at d_i,

P_t = spot price of underlying equity,

K = exercise price of option,

r_f = continuously compounded risk-free interest rate,

$T-t$ = time to maturity of option in years.

[10] For further details on the BS model, see Black, Fischer and Myron Scholes (1973), The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* Vol. 81, 637-654.

By plotting the implied volatility against the maturity of the associated option, a term structure of the implied volatilities can be inferred. Based on its design, an implied volatility represents an estimate for the average volatility over the life of the particular option instead of the instantaneous, overnight volatility. This means that if price quotes are only available for longer-term options, the volatility surface need to extrapolated to the near term.

The implied volatility approach is not always so robust. One potential objection to the use of the approach is that the BS model has shown to be inconsistent with stochastic volatilities. Recent empirical findings on the effect of stochastic volatilities, however, indicate that the BS model performs well for short-term at-the-money options. For other types of options, such as deep out-of-the-money options, the model may be less appropriate, creating discrepancies in implied volatilities. These discrepancies form a feature generally called volatility smile. More precisely, the volatility smile refers to the situation when options with the same characteristics except the strike price imply different volatilities. (Jorion, 2000)

One of the major attractions of the implied volatility approach is that market participants have sufficient confidence in them by betting real money on them. Since options' prices reflect the market consensus about future volatilities, there are sound reasons to believe that the implied volatilities are superior to estimates based on historical data. Jorion (2000) argues that the empirical evidence indeed indicates the superiority of options data, and he strongly supports the use of implied volatilities in VaR calculations, whenever it's possible.

Unfortunately, the implied volatility approach doesn't suit so well for the variance estimation in this study. Since variance estimates are needed for altogether 138 different equities and indices for 1515 different days, data would be needed at least on 209070 option prices and other specific option characteristics, such as exercise prices. Availability of such comprehensive data is clearly a serious problem. Furthermore, in order to get daily and weekly variance estimates from longer-term options, extrapolation should be carried out in a non-automated way. For these reasons, the use of implied volatilities doesn't seem to be worthwhile in this study. Thus, variance estimation has to be carried out with one of the earlier described methods that use historical return series data.

3.4.5. Comparison of methods and selection of appropriate method

As pointed out in the previous chapter, implied volatility approach is ill-suited for this study leaving the three historical data-based methods left for selection. The basic features of the three methods are presented earlier in their respective chapters, and the purposes of this chapter are to further point out and compare important characteristics of the methods and select an appropriate variance estimation method for the study.

Although Chapter 2.2 already presents some research findings concerning the VaR measure accuracy performance of different variance estimation methods, none of those studies so thoroughly compare different methods, as Alexander and Leigh (1997) do. Therefore, their study and findings are worth interest, and are presented here.

Alexander and Leigh (1997) compare the performance of the MA, GARCH(1,1) and EWMA variance estimation methods. They evaluate the methods both statistically, using likelihood and root mean squared error (RMSE) evaluation procedures, and operationally, using the evaluation procedure proposed by the Bank for International Settlements (BIS). In statistical evaluation, they assess how well the methods estimate the center of return distributions, and find that the EWMA method is generally the most accurate method for five-, ten-, and twenty-five-day returns; in daily returns the results are very mixed.

In VaR environment, it's however the worst returns, not the most center ones, that should be predicted as accurately as possible. Thus, in their operational evaluation, Alexander and Leigh (1997) examine the methods' ability to predict the lower percentiles of return distributions. Their findings indicate that the MA and GARCH(1,1) methods generally produce variance estimates, which will result in accurate 99% VaR measures for equities and currencies. In contrast, they find that EWMA method produces in general too low variance estimates, and applying this method in VaR calculations could lead to an unacceptably high number of outliers.

The findings of Alexander and Leigh (1997) shouldn't be taken as the final truth of the competencies of the three methods, especially since their test period is relatively short, only 200 days, and there exist various other testing procedures that could be used to evaluate the variance estimation methods. Furthermore, some of the studies presented in Chapter 2.2, such as studies of Mahoney (1995) and Hendricks (1996), find that EWMA method produces accurate 95% VaR measures. Nevertheless, Alexander and Leigh's study adds valuable information concerning the selection of an appropriate variance estimation method. The earlier discussion in this chapter more or less discourages application of the MA method. But despite of its disadvantages, such as the ghosting feature, according to Alexander and Leigh the MA method produces accurate variance estimates to VaR calculations, while the EWMA method is not necessarily as accurate and effective as the earlier discussion in this chapter indicates.

An especially important issue concerning variance estimation in this study is time aggregation. Particularly, how well the three different methods can be used to extrapolate daily variance estimates to longer horizons. Obviously, one could estimate weekly variances by using weekly returns data, monthly variances using monthly returns data, and so on. However, this isn't in general desirable. Using higher-frequency data is generally more efficient because it uses more available information (Jorion, 2000).

Since this study uses geometric returns, which are additive as shown earlier in Chapter 3.3 in Formula (27), the time aggregation of variances can be done in a convenient way, when three specific simplifying assumptions are made, all of which are generally considered relatively reasonable in research. Firstly, it's assumed that returns are uncorrelated over successive time periods. This assumption is consistent with informationally-efficient markets, where the prevailing market price reflects all relevant information about a particular asset. If this holds, all price changes must result from relevant news that, by definition, cannot be anticipated and therefore must be uncorrelated over time; in other words, prices follow a random walk. Thus, the covariance between successive periods' returns $\mathrm{cov}(R_t, R_{t-1})$ must be zero.

The second assumption is that returns are identically distributed over time. So, it's assumed that the following holds for variances: $\mathrm{var}(R_t) = \mathrm{var}(R_{t-1}) = \mathrm{var}(R) = \sigma^2$. The first and second assumptions form together the general i.i.d. assumption of returns; i.i.d. stands for independently and identically distributed.

The third and last assumption, or rather a condition, is that the positions are constant over the extended period. Since no trading occurs for the portfolios tested in the study, the previously stated condition holds. Based on the i.i.d. assumption of returns, calculation of the two-period variance is a straightforward task, and can be done using Formula (35) below.

$$\sigma_{t,2}^2 = \mathrm{var}(R_t + R_{t-1}) = \mathrm{var}(R_t) + \mathrm{var}(R_{t-1}) + 2\,\mathrm{cov}(R_t, R_{t-1}) = 2\sigma^2 \qquad (35)$$

Thus, the two-period variance is simply the periodic variance multiplied by two. Since covariances are expected to be zero across different periods, this framework can be easily developed further to cover as many periods as needed. For instance, by holding the i.i.d. assumptions, a weekly return variance is five times a daily return variance; the number five is applied for weekly estimates instead of seven, since there are generally five trading days in a week. A weekly standard deviation or volatility is respectively a square root of five multiplied by a daily standard deviation. Under i.i.d. the possibility to calculate multiple-period volatilities by multiplying the periodic volatility with the square root of multiple is generally called the square root of time rule.

As the MA method makes an implicit assumption that the true volatility is constant over time, it thereby also assumes that the future return variances are identical within periods of similar lengths. If it's further assumed that returns are independent of each other, the square root of time rule can be applied to the MA volatility estimates without any contradictions. Unfortunately, the use of the rule becomes more complicated, when the variances are estimated either with the GARCH(1,1) or EWMA method.

Assuming the returns are uncorrelated across days, the GARCH(1,1) model can be used to extrapolate next day's variance estimate to longer horizons in a consistent fashion. Given that the GARCH(1,1) model is stationary and the persistence parameter $(a_1 + b)$ is less than one, the extrapolation for an L-day variance can be made using the next page's Formula (36). (Jorion, 2000)

$$\sigma_{t,L}^2 = \frac{a_0}{1-(a_1+b)}\left[(L-1)-(a_1+b)\frac{1-(a_1+b)^{L-1}}{1-(a_1+b)}\right] + \frac{1-(a_1+b)^L}{1-(a_1+b)}\sigma_t^2 \tag{36}$$

where

σ_t^2 = GARCH(1,1) estimate for next day's variance that is found through Formula (30).

Formula (36) shows that the extrapolation is a complicated function of the variance process and the initial situation σ_t^2. Consequently, the simple square root of time rule can't be used in this case, because according to the stationary GARCH(1,1) process the returns are not identically distributed over time. However, there do exists one special case in the GARCH(1,1) environment when the square root of time rule is valid: If the initial position happens to be equal to the long-term average value σ^2, which is given by Formula (31), the extrapolation Formula (36) yields $\sigma_{t,L}^2 = L\sigma^2$ that is consistent with the square root of time rule. Generally speaking, if the initial position is greater than the long-term average value, the square root of time rule will overestimate risk. On the other, if the initial position is less than the long-term average value, the rule will underestimate risk. (Jorion, 2000)

If the EWMA method is applied in the variance estimation, the next period's variance estimate σ_t^2, which is obtained by Formula (32), can be used to determine the following period's variance estimate σ_{t+1}^2 in a straightforward way. Since the EWMA method, as each GARCH model, assumes that the variance of returns follows a predictable process, it can be derived that

$$\sigma_{t+1}^2 = E_{t-1}(R_{t+1}^2) = E_{t-1}\left[\lambda\sigma_t^2 + (1-\lambda)R_t^2\right] = \lambda\sigma_t^2 + (1-\lambda)\sigma_t^2 = \sigma_t^2 . \tag{37}$$

Thus, in contrast to the stationary GARCH(1,1) model, the daily variances estimated with the EWMA method are expected to be identical over time. This comes as no surprise, since the persistence parameter $(a_1 + b)$ is one for the EWMA method. So, the effect of the possible shock in R_{t-1}^2 will fully remain in all future extrapolated variance estimates. In other words, the EWMA method allows no mean reversion in the extrapolation. However, Jorion (2000) states that mean reversion is usually observed in longer-term variance estimates. This means that the validity of the extrapolation in the EWMA method is very questionable. Especially, if the shock is large or the length of the extrapolation period is long, the use of extrapolation may yield seriously inaccurate estimates. Therefore, even though the theoretical basis exists for the square root of time rule in the EWMA environment, if the assumption is made on independent returns, the use of the rule doesn't seem to be sane.

Due to the extrapolation problems in the EWMA approach, RiskMetrics formulates a separate formula for monthly variance estimates beside the original EWMA Formula (32) that they apply for daily variance estimates. RiskMetrics defines the month as 25 trading days, and their monthly variance estimates are found through Formula (38), which is shown below. (RiskMetrics – Technical Document, 1996)

$$\sigma_{t,25}^2 = \lambda\sigma_{t-1,25}^2 + (1-\lambda)s_{t-1}^2 \tag{38}$$

where

$$s_{t-1}^2 = \sum_{k=1}^{25} R_{t-k}^2 .$$

Comparing (32) and (38) it can be seen that there's merely one essential difference between these two formulas or models. In other words, Formula (38) redefines R_{t-1}^2 of Formula (32) as the 25-day moving variance estimator s_{t-1}^2. However, this is not free from undesired effects, as it creates in practice ghosting. Thus, the MA method is not the only method that suffers from ghosting, when longer than daily estimates are concerned.

The ghosting feature has slightly different implications in the monthly EWMA method than it has in the MA method. Alexander and Leigh (1997) point out that the monthly variance estimates obtained through Formula (38) achieve their maximum 25 days after a major market shock. They also explain why this happens: From (38) it can be reasoned that $\sigma_{t,25}^2 > \sigma_{t-1,25}^2 \Leftrightarrow s_{t-1}^2 > \sigma_{t-1,25}^2$. This means that the monthly EWMA variance estimate will continue to rise while the daily MA variance estimated from the past 25 days remains artificially high during the ghost feature. Exactly 25 days after the market shock that caused the feature, s_{t-1}^2 will drop dramatically, so the maximum value of $\sigma_{t,25}^2$ will occur at this point.

Based on the considerations presented in the chapter, the MA seems to be the most appropriate variance estimation method for the purposes of this study. Since a large number of daily variance estimates, altogether 209070 estimates, is needed in the study, the estimation method should be easy to implement. MA variance estimates are easily found, since required data is readily available and the method involves no parameter optimisation.

Because estimates are needed also for 208518 weekly variances, the estimation method should be able to use effective extrapolation measures to daily estimates; the fact is that it's generally more efficient to use high-frequency daily data instead of weekly data. When the MA method is applied, and the i.i.d. assumption is made on returns, an estimate for particular weekly variance is obtained with ease by multiplying the prevailing daily variance estimate by five.

In the case of the GARCH(1,1) or EWMA method, the use of the square root of time rule is problematic. Also, the theoretical extrapolation in the GARCH(1,1) environment through Formula (36) could turn out to be problematic, since the optimal parameter values may change over time. On the other hand, due to the EWMA method's persistence of one, a special model, similar to RiskMetrics' monthly EWMA model, should be designed for weekly variance estimates. This would require an in-depth study on the weekly estimates' optimal decay factor λ.

Finally, the findings of Alexander and Leigh (1997) indicate that the MA estimates work accurately when VaR measures are concerned. Therefore, there doesn't seem to be any evident reason to use a more sophisticated estimation method for variances instead of the traditional MA method. Whether or not the application of the MA method for variance estimation lessens the accuracy of the VaR models tested can be examined in further studies.

In the study the variances are estimated from periods of different lengths. This is due to the fact that there isn't any universal optimal estimation period length, which should be applied in the MA method. As the discussion in Chapter 3.4.1 points out, longer estimation periods reduce the ghosting feature, whereas shorter estimation periods may more effectively gauge the underlying evolution in variance. The three estimation periods lengths applied in the study are 250, 500 and 1000 days. Since many financial institutions act under the BIS guidelines, which require that at least one year of historical data must be used in the variance estimation, the estimation period lengths of less than 250 days are not applied, though they might improve the estimation accuracy (Basle Committee on Banking Supervision, 1996). Weekly standard deviation or volatility estimates are found applying the square root of time rule and multiplying the daily volatility estimates by square root of five.

3.5. Estimation of covariances

Since three of the study's five models, in particular Traditional model (7), Sector-beta model (23) and Diagonal sector-beta model (24), include a covariance matrix, also covariance estimates are needed in addition to the variance estimates. There exist very similar estimation methods for covariances as they do for variances. Unfortunately, the practicality scale of the methods is even wider than it is among the variance estimation methods.

Again, the simplest method is the moving average (MA). It has all the same basic properties as the MA variance estimator, which result from weighting each past return observation equally. The covariance between two return series, 1 and 2, is estimated applying the MA method as follows

$$\sigma_{12,t} = \frac{\sum_{i=1}^{M}(R_{1,t-i} - \overline{R}_1)(R_{2,t-i} - \overline{R}_2)}{M-1}. \tag{39}$$

GARCH estimation can be applied to covariances, too. This can be done by extending the univariate framework, which is applied to single variances, to the multivariate one.[11] However, application of the multivariate GARCH procedures becomes rapidly reluctant, as

[11] For specific information on the designs of the multivariate models, see for example Campbell, Lo and MacKinlay (1997).

the portfolio size increases. The problem is that the number of parameters, which need to be estimated, grows exponentially with the number of series. With two equities in the portfolio, for instance, estimates are required for nine parameters; three a_0, a_1 and b_1 parameters for each of the three covariance terms (Jorion, 2000). When the portfolio size increases, there comes a point where there are so many parameters that they cannot be estimated. Moreover, multivariate GARCH methods can with ease produce correlation estimates that are outside the permissible range [-1 , 1]. (Dowd, 1998)

The EWMA approach can be used also for covariance estimation. The following EWMA covariance estimation Formula (40) looks quite identical with the variance Formula (32)

$$\sigma_{12,t} = \lambda\sigma_{12,t-1} + (1-\lambda)R_{1,t-1}R_{2,t-1}\,. \tag{40}$$

By using a same value for decay factor λ in all EWMA estimates for variances and covariances with identical horizons, the parameter estimation problems of the multivariate GARCH framework can effectively be avoided. This is actually what RiskMetrics does. In their system, similarly to the variance estimation, the decay factor λ takes a value of 0.94 for daily and a value of 0.97 for monthly covariance estimates (RiskMetrics - Technical Document, 1996). Consequently, the EWMA method is guaranteed to produce a correlation coefficient that is in the permissible range (Dowd, 1998).

Options data can be used not only for implied volatilities but also for implied covariances. There exist options, for example so-called quanto options, whose payoffs are dependent on the prices of two underlying variables. The valuation formula of such an option involves also the covariance between the two return series. After recovering the two implied volatilities from common options whose payoffs depend only on single underlying variables, the price of a quanto option can be used to infer an estimation for the covariance. (Jorion, 2000)

Although the implied covariances are presumably rather efficient estimates, since they are totally based on future expectations, their application raises problems. Firstly, the availability of data on such a special kinds of options is limited, at least currently. Secondly, the covariances should be recovered from the options' prices in a non-automated way, which makes this approach fairly unpractical.

The previous discussion on covariance estimation methods clearly indicates that only two true alternatives exist for the estimation, particularly the MA and EWMA methods. Jorion (2000) compares the correlation estimates achieved through the MA and EWMA methods. Formally, the correlation estimate $\rho_{12,t}$ is obtained from the estimated variances and covariances as follows

$$\rho_{12,t} = \frac{\sigma_{12,t}}{\sqrt{\sigma_{1,t}^2 \sigma_{2,t}^2}}. \tag{41}$$

In his comparison, Jorion (2000) uses two estimation period lengths for the MA method: 20 and 60 days. For the EWMA method, he uses the decay factor λ value of 0.94, the same as RiskMetrics uses for daily estimates. Jorion studies the correlation between the \$/BP exchange rate and the \$/DM rate over the 1990-1994 period. He finds that the correlation estimates of the EWMA method do not diverge significantly from the MA estimates; generally the EWMA estimates lie between the MA(20) and MA(60) estimates. Thus, Jorion's findings indicate similarity of the MA and EWMA methods in the correlation estimation; it should, however, be noted that the relatively short estimation period lengths in the MA method weakens the competency to generalise the findings.

Jorion (2000) also points out that since the weights of the past observations decay quite rapidly in the EWMA method and the number of effective observations is therefore small, the application of the method could easily result in a covariance matrix, which fails to be positive semi-definite. Although this problem is obviously more essential the more the portfolio includes assets, it nevertheless decreases attractiveness of the EWMA approach.

Moreover, as variances are estimated in the study using the MA method (28), for consistency reasons and for the theoretically valid use of the square root of time rule on daily portfolio standard deviations, this study uses the MA method (39) for covariance estimation. The same three estimation period lengths, particularly 250, 500 and 1000 days, are also applied in the covariance estimation. Daily portfolio standard deviations are multiplied by square root of five to get estimates of weekly portfolio standard deviations; thus, the potential daily covariance terms are thereby implicitly multiplied by five.

3.6. Estimation of betas

The estimation of betas β_i in the market models (8) and (18) is made by applying the ordinary least squares (OLS) estimation procedure. OLS is based on finding the parameter values α_i and β_i, which minimise the sum of the squared residuals $\sum_{j=1}^{M}\varepsilon_j^2$ in the sample period. Under general conditions OLS is considered to be a consistent estimation procedure for the market model parameters. Furthermore, assuming that asset returns are jointly multivariate normal and i.i.d. through time, OLS is efficient. (Campbell, Lo and MacKinlay, 1997)

The OLS estimator for the equity *i*'s market beta for day *t* is

$$\beta_{i,t} = \frac{\sum_{j=1}^{M}(R_{m,t-j} - \overline{R}_m)(R_{i,t-j} - \overline{R}_i)}{\sum_{j=1}^{M}(R_{m,t-j} - \overline{R}_m)^2} = \rho_{im,t}\frac{\sigma_{i,t}}{\sigma_{m,t}}. \tag{42}$$

Individual equity's sector beta is calculated similarly to the market beta by replacing the market index returns R_m :s with particular sector index returns R_{Si} :s in Formula (42). In the study, the beta estimates are calculated from the daily return data for both the daily and weekly VaR measures. Furthermore, the beta estimates are updated each day, and consistently with the variance and covariance estimation, the same three estimation period lengths, to be precise 250, 500 and 1000 days, are applied in the beta estimation.

3.7. Tests and corresponding hypotheses

This study uses three kinds of statistical tests in attaining the results. More specifically, the tests are applied in order to find out how well certain assumptions concerning the portfolio returns hold and how accurate the VaR models studied seem to be from the statistical viewpoint. All the tests and their respective hypotheses are presented in the following three subchapters.

3.7.1. *t*-test for zero mean portfolio return

Because relative VaR measures are calculated in the study, the calculation of the actual measures doesn't involve portfolio's expected return μ_p. However, when the VaR measures are backtested and compared to the observed portfolio outcomes, the portfolio's expected return has to be taken into account. By its design a relative VaR measure indicates the loss relative to the expected outcome. Thus, to ensure properly specified backtesting framework, the expected portfolio returns, provided that they deviate from zero, have to be subtracted from the observed portfolio returns.

Especially for the daily equity returns, it's generally assumed that the expected return is zero. If this holds, no amendments have to be made to the observed portfolio returns when the VaR measures are becktested. Therefore, statistical testing is applied in the study to find out whether the assumption of zero expected return is acceptable or not.

Mean portfolio return of a sample $\overline{R}_p$ is an unbiased, efficient and consistent estimator of the expected portfolio return. In general, unbiasedness means that the estimator's average value is equal to the population parameter being estimated. On the other hand, efficiency requires that, for a given sample size, the standard error of the estimator's numerical value is as small as possible. Furthermore, a sample estimator is said to be consistent if its value approaches the value of the population parameter being tested as the sample size increases. (Fleming and Nellis, 1994)

Due to its favourable properties, the sample mean return $\overline{R}_p$ is used through a *t*-test to test the appropriateness of the zero expected portfolio return. The null hypothesis that the expected portfolio return equals zero is tested for each of the three portfolios, for both daily and weekly returns, on the VaR measure test period 1995-2000. Since it's hard to imagine negative expected returns on the portfolio, the alternative hypothesis is that the expected portfolio return is greater than zero.

As the standard error of the mean portfolio return $\sigma_p/\sqrt{s}$ is unknown and has to be estimated from the sample of size *s*, the test statistic *t* doesn't follow the standard normal distribution but rather *t* distribution with *s*-1 degrees of freedom.[12] The hypotheses (43) as well as the test statistic (44) of the one-tailed *t*-test applied in the study are formally

$$H_0: \mu_p = 0 \quad H_1: \mu_p > 0 \tag{43}$$

$$t = \frac{\overline{R}_p - 0}{\sigma_{R_p}/\sqrt{s}} = \frac{\overline{R}_p}{\sigma_{R_p}/\sqrt{s}} \sim t(s-1). \tag{44}$$

The null hypothesis of zero expected portfolio return is then rejected, if the test statistic gets a larger than critical value. The proper critical value for the *t*-test depends on the sample size, or degrees of freedom, and the chosen significance level. The larger the sample size the lower the critical value. The significance level indicates the probability that the valid null hypothesis is incorrectly rejected; this is generally called the probability of committing a type 1 error. So, the lower the significance level is, the higher the critical value. In the study, the critical values for the *t*-tests are obtained directly through a computer program; they could be obtained also from standard statistical tables.

3.7.2. Anderson-Darling test for normality of portfolio returns

All the VaR models tested in the study are based on the assumption that the portfolio returns are normally distributed. If they truly are normally distributed, then the potential weaknesses of the VaR models tested are due to inaccurate portfolio standard deviation estimates. On the other hand, if the portfolio returns don't seem to follow the normal distribution, the potential inaccuracies of the VaR models tested can arise from poor

---

12 However, the differences between the percentage points of the standard normal and *t* distributions are rather infinitesimal when the sample size *s* is over 30. Since sample sizes of over 1500 are used in the study for the testing, the standard normal distribution could be used as well for the test statistic *t* without any significant violations.

standard deviation estimates and also from the misspecification in Formula (1). Therefore, it's worth testing the normality of the portfolio returns.

It's important to point out here that the Anderson-Darling test applied tests whether the unconditional distribution of the portfolio returns is distributed normally, while the variance-covariance approach actually assumes that the conditional distribution of the portfolio return is normally distributed. In other words, the test could easily reject the hypothesis of normality, even though the portfolio returns at each point of time in reality would come from the normal distribution.

Nevertheless, in accordance with the i.i.d. assumption, it's expected in the study that the parameters of the normal distribution, that is the expected portfolio return μ_p and the portfolio standard deviation σ_p, remain sufficient stable during the test 1995-2000 period so that the results of the Anderson-Darling tests are fairy reliable. In other words, a rough assumption is made in the study that what holds for the unconditional portfolio return distributions holds also for the conditional distributions.

The Anderson-Darling test applied in the study to the normal distribution is a kind of goodness-of-fit test; in other words, it tests does the hypothesised normal distribution fit to the observed distribution. While the Anderson-Darling test is on the whole a powerful test to the normal distribution, it's designed to be more sensitive to the discrepancies between the hypothesised normal and empirical distributions in the tails of the distribution (Sinclair and Spurr, 1988). Thus, the test suits particularly well for the purposes of this study, which is keen on the lower tails of the distribution.

D'Agostino and Stephens (1986) as well recommend the use of the Anderson-Darling test. They also provide a six-step procedure for performing the test, when both the expected return μ_p and standard deviation σ_p are unknown.[13] The hypotheses of the test are

$$H_0: \textit{Portfolio returns are normal} \qquad H_1: \textit{Portfolio returns are not normal.} \tag{45}$$

[13] The application of other normality tests, such as Kolmogorov-Smirnov, Kuiper, Cramer von Mises and Watson tests, is also described in D'Agostino and Stephens (1986).

The actual test is carried out through the following six steps:[14]

1. Portfolio returns in the sample of size s are arranged in ascending order,

$$R_{p,1} \leq \cdots \leq R_{p,s}. \tag{46}$$

2. Standardised values Y_i :s are calculated,

$$Y_i = \frac{R_{p,i} - \overline{R}_p}{\hat{\sigma}_p}. \tag{47}$$

3. Cumulative probabilities X_i :s are calculated by the standard normal distribution,

$$X_i = N(Y_i) = \int_{-\infty}^{Y_i} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt. \tag{48}$$

4. The Anderson-Darling test statistic A^2 is computed,

$$A^2 = -s - \sum_{i=1}^{s} \frac{(2i-1)}{s} [\ln X_i + \ln(1 - X_{s+1-i})]. \tag{49}$$

5. The modified test statistic A^* is computed,

$$A^* = A^2 \left(1.0 + \frac{0.75}{s} + \frac{2.25}{s^2} \right). \tag{50}$$

6. The null hypothesis of normality is rejected if the modified statistic A^* exceeds 0.631, 0.752, 0.873, 1.035 and 1.159 at the levels of significance 0.10, 0.05, 0.025, 0.01 and 0.005, respectively. (D'Agostino and Stephens, 1986)

[14] This procedure is valid for samples of size $s \geq 8$.

Similarly to the t-test, the previously described Anderson-Darling test is applied on each of the study's three equity portfolios, for both daily and weekly returns, on the VaR measure test period 1995-2000.

3.7.3. Kupiec tests for exceeded VaR measures

In order to test the accuracy of the study's VaR models, the observed portfolio return outcomes are compared to the calculated VaR measures on the test period 1995-2000. If the observed portfolio loss is larger than the particular VaR measure, the term exception is used for this event. Ideally, there would be 5% exceptions for the 95% VaR measures and respectively 1% exceptions for the 99% VaR measures on the test period. Obviously, such exact percentages are very unlikely.[15] The question is then, when the departures from the ideal failure rates are so large that they can't be simply explained by bad luck. In other words, when application of a particular VaR model is inappropriate.

Unfortunately, there aren't currently any robust method to determine the accuracy of a VaR model. Haas (2001) considers the existing backtesting methods, such as two Kupiec tests, Lopez' magnitude loss function and Crnkovic and Drachman (CD) test, finding them more or less weak and inappropriate. For the improved backtesting, Haas introduces a mixed Kupiec test and two modifications of the CD test, which he calls a scaled CD test and a weighted scaled CD test. He studies the introduced three backtesting methods empirically on various test periods for equity VaR measures, which are calculated applying the variance-covariance approach and the historical simulation with different estimation period lengths. Haas finds that all the three backtesting methods judge the quality of a VaR model reasonably well, the mixed Kupiec test performing most effectively. This is mainly due to the mixed Kupiec test's ability to identify dependencies in exceptions, which the CD tests are unable to do.

Based on the findings of Haas (2001), this study uses the mixed Kupiec test and its components for backtesting the VaR models. Failure rate tests are in general criticised for

[15] In fact, since this study has 1515 portfolio return observations for daily VaR measures and respectively 1511 observations for weekly VaR measures, which both are not divisibly evenly by 100, the exact exception percentages of 5 and 1 are impossible.

their low statistical power when only a couple of hundreds observations are available for backtesting. However, because this study has a relatively large number of observations on the test period, more precisely over 1500 observations, the use of the failure rate tests should be rather powerful. Foundations and characteristics of the Kupiec tests are presented next.

Defining x as a number of exceptions on the test period, a VaR model's failure rate x/s, which is as an estimator $\hat{p}$, should converge to 1-VaR model confidence level as the sample size increases, if the model provides accurate coverage. Under the null that a particular VaR model is accurate, the hypotheses of the failure rate test can be written formally as follows[16]

$$H_0: \ p = 1\text{-}VaR \ \textit{model confidence level} \qquad\qquad H_1: \ p \neq 1\text{-}VaR \ \textit{model confidence level.} \tag{51}$$

The setup for the failure rate test is the well-known testing framework for a sequence of successes and failures, also called Bernoulli trials. Thus, the number of exceptions x follows a binomial probability distribution, as shown formally below in (52). Based on the characteristics of the binomial probability distribution, x has an expected value $E(x) = ps$ and a variance $\text{var}(x) = p(1-p)s$. (Jorion, 2000)

$$f(x) = \binom{s}{x} p^x (1-p)^{s-x}. \tag{52}$$

The binomial distribution formula, as it stands above, could be used to set critical values for the failure rates. For instance, Jorion (2000) uses for the failure rate x/s a critical value of 2%, his sample size being 250 and VaR confidence level 99%. In this case, Formula (52) implies that there's a 10.8% probability of committing a type 1 error and rejecting a correct model. Moreover, as Jorion shows with an incorrect model indicating a true p of 3%, application of Formula (52) for critical values imply relatively high, easily over 10%,

16 There shouldn't naturally be too many exceptions but it's also inappropriate to have too low number of exceptions and too conservative VaR model, which is the reason why generally a two-tailed test is applied.

probabilities for accepting an incorrect model. In other words, this framework creates a very high type 2 error.

Kupiec (1995) develops a failure rate test that is also based on the binomial distribution of exceptions. However, his test is more powerful than the previously described basic framework and creates a lower type 2 error.[17] Kupiec's failure rate test is applied to the same hypotheses as shown earlier in (51). The test is so called likelihood ratio (LR) test, and its corresponding LR test statistic is defined as

$$LR_{uc} = -2\ln\left(\frac{p^x(1-p)^{s-x}}{(x/s)^x\left[1-(x/s)\right]^{s-x}}\right). \tag{53}$$

According to Kupiec (1995) the test statistic is asymptotically distributed chi-square χ^2 with one degree of freedom under the null hypothesis. Therefore, the critical values for the test can be readily obtained, for example, from the statistical tables reporting the percentage points of the $\chi^2(1)$ distribution.

Exceptions may well be clustered over time, which should invalidate a VaR model. Thus, conditional coverage of a VaR model should be tested. The above test statistic LR_{uc} measures only the unconditional coverage of the model, since it ignores time variation in the data. The independence of the exceptions should also be considered in order to test the conditional coverage. Christoffersen (1998) develops an independence test, which measures whether the probability of an exception on a particular day is dependent on the previous day's outcome. However, Haas (2001) argues that Christoffersen's test is too weak to deliver feasible results and he introduces an improved test for independence of exceptions, which is build on the time until first failure (tuff) test developed by Kupiec (1995).

The tuff test measures the time until the first exception occurs for a particular VaR model. The binomial distribution of exceptions implies that an exception is expected to occur every $1/p$ time periods; obviously, these time periods are days for daily VaR measures and

[17] Kupiec (1995) argues that based on the Neyman-Pearson lemma, his test is the most powerful among its class.

respectively weeks for weekly VaR measures. Based on this, Kupiec (1995) presents the following LR test statistic for his tuff test

$$LR_{tuff} = -2\ln\left(\frac{p(1-p)^{v-1}}{(1/v)[1-(1/v)]^{v-1}}\right) \tag{54}$$

where

v = number of time periods until the first exception occurs.

The tuff test can be used to test the hypotheses (51), and the test statistic LR_{tuff} is again asymptotically $\chi^2(1)$ distributed. Though the test is according to Kupiec (1995) very weak in discriminating among alternative hypotheses, Haas (2001) finds it useful for his independence test and forms a test statistic for the time between two successive exception

$$LR_i = -2\ln\left(\frac{p(1-p)^{v_i-1}}{(1/v_i)[1-(1/v_i)]^{v_i-1}}\right) \tag{55}$$

where

v_i = number of time periods between exceptions *i* and *i*-1.

Haas' (2001) test statistic LR_i for time between failures can be calculated for every exception on the test period, except for the first exception. By adding the tuff test statistic LR_{tuff} of the first exception, a total of x, which is the number of exceptions, test statistics is received. Assuming that the exceptions are independent from each other, the test statistics are independent as well and can be summed. Since χ^2 distribution is also additive, the critical values can be added, too. As a result, a test for independence is achieved. Thus, the hypotheses of Haas' test and the corresponding test statistic are

$$H_0: \textit{Exceptions are serially independent} \quad H_1: \textit{Exceptions aren't serially independent} \tag{56}$$

$$LR_{ind} = \sum_{i=2}^{x}\left(-2\ln\left(\frac{p(1-p)^{v_i-1}}{(1/v_i)[1-(1/v_i)]^{v_i-1}}\right)\right) - 2\ln\left(\frac{p(1-p)^{v-1}}{(1/v)[1-(1/v)]^{v-1}}\right). \tag{57}$$

The test statistic LR_{ind} is asymptotically chi-square distributed with x degrees of freedom, that is $\chi^2(x)$. By combining the test statistics of unconditional coverage LR_{uc} from (53) and independence LR_{ind} from (57), Haas (2001) presents a mixed test statistic LR_{cc} shown below in (58), which can be used to test the conditional coverage of a VaR model; hypotheses of this test are similar to the hypotheses in (51), p now representing the conditional probability of an exception.

$$LR_{cc} = LR_{uc} + LR_{ind} \tag{58}$$

Since the test statistic of unconditional coverage LR_{uc} is independent of all the time between failures test statistics LR_i:s, the test statistic of conditional coverage LR_{cc} is, in accordance with the other test statistics, asymptotically $\chi^2(x+1)$ distributed. This study calculates both the test statistic LR_{cc} as well as separately its components LR_{uc} and LR_{ind}, and tests their respective hypotheses for each of the VaR models. The components of the conditional coverage test statistic LR_{cc} are calculated separately, so that the reason for potential rejection of the conditional coverage hypothesis can be found; in other words, by calculating the test statistics LR_{uc} and LR_{ind} also separately, it can be found out whether the conditional coverage hypothesis of a VaR model is possibly rejected due to a weak general coverage or dependence of the exceptions, or perhaps due to the both.

One central issue concerning backtesting is the significance level of the tests. Backtesting involves balancing type 1 and 2 errors. Based on the examination of Kupiec (1995), in interpreting the likelihood ratio test statistics, the 5% significance level seems to be preferable to the other common significance level of 1%. This is because application of the 1% significance level presumably leads to unacceptably high type 2 errors, which more than negates the 4% advancement in type 1 error compared to the 5% significance level. In other words, the study's VaR model accuracy hypotheses that aren't rejected at the significance level of 1% but are significant at the 5% level should be considered with special caution.

4. Data and portfolio characteristics

4.1. Data

The beta mapping-based VaR models tested in the study are derived from the market models. Since the market models (8) and (18) relate the equity return to the return of the market or sector index on the same day, the equities included in this study should be liquid. If the equities were not liquid, the OLS estimation would produce downward biased betas since the equity price changes due to the general market or sector changes would be lagged, at least partly; in other words, the true dependence of the illiquid equities' returns on the market or sectors can't be found out by simply comparing the changes on one particular day.[18] With the downward biased betas, the beta mapping-based VaR models would clearly produce too low VaR measures because betas are the key factors determining volatilities in these models.

Mainly due to these liquidity reasons, this study uses data from the U.S. equity markets. In the financial markets, S&P 500 is one of the most popular general market indices and it's used in this study. In addition, S&P 500 sector indices are used for the study's sector mapping VaR models. Constituents of the S&P 500 index represent equities that are actively traded, and they should thus be liquid (S&P 500 Index Methodology, 2001). Therefore, equities included in the S&P 500 index are used in composing the study's three portfolios.

The data is collected from Datastream database. For the return calculations, data is needed from years 1991-2000 on a business daily basis. Price index values, which are used in the study to calculate the returns of the non-dividend paying market and sector indices, are readily available from Datastream for the S&P 500 general market index. Unfortunately, with the sector indices the situation is slightly problematic, and price index values are collected only for eight sector indices out of eleven. Datastream does report price index values for all eleven S&P 500 sector indices, but for three of the sectors, particularly

[18] Empirical evidence shows that when beta estimation is made applying methods which take into account of thin trading, the beta estimates of the illiquid equities are approximately 10 to 20% larger than their corresponding OLS betas. For thin trading-adjusted beta estimation methods, see for instance Scholes, M. and J. Williams (1977), Estimating Betas from Nonsynchronous Data, *Journal of Financial Economics* Vol. 5, 309-328.

Communication Services, Consumer Cyclicals and Consumer Staples, the values are reported only from 26th of June 1996 onwards. Thus, taking into account the required data period starting from 1991, the three mentioned sectors have to be excluded from the study's perspective. In other words, only equities representing the remaining eight sectors are eligible for the study, and the portfolios are composed of these equities.[19]

At first, for practical reasons, a sample of 155 equities is randomly picked from the eligible S&P 500 constituents.[20] Then, the study's three portfolios are composed by further taking random sub-samples of 30, 50 and 100 equities from the initial sample. As a result, the three portfolios include altogether 129 different equities. For each selected equity a total return index (TRI) series covering years 1991-2000 is collected from Datastream. TRI_t is calculated in Datastream as follows

$$TRI_t = TRI_{t-1} \frac{P_t + D_t}{P_{t-1}}. \tag{59}$$

As (59) clearly indicates, by taking natural logarithm of two successive TRI_t:s, an equivalent calculation to (26) is obtained. Therefore, the application of total return indices for the equity return calculation is justified. In addition to the total return indices, for each of the study's equities, the market capitalisation at the beginning of the year 1995 is collected from Datastream for determination of equity and sector weights, as described in Chapter 3.2. Only the initial market capitalisations at the beginning of the test period are required since the calculated equity returns are used to determine the subsequent weights in the portfolios.

4.2. Portfolio characteristics

As pointed out earlier, the compositions of the study's small, medium and large portfolios are fixed during the test period 1995-2000, but the weights of the individual equities and

[19] Naturally, equities of the three excluded sectors could be used in testing Traditional model (7) and the two VaR models (16) and (17) that apply mapping onto the market, but using different data on different VaR models would make the comparisons between all the five models more problematic.

[20] Constituents of the S&P 500 index vary through time. This study collects the data from the S&P 500 constituents of 31st of August 2001. Obviously, it would be more appropriate to use constituents that more effectively reflect the test period 1995-2000, but this kind of data is unfortunately not available.

sectors do change according to the changes in the market prices. The value development of S&P 500 and the study's portfolios during the test period is described in Figure 1 below.

Figure 1.
S&P 500 and portfolio value development during 1995-2000

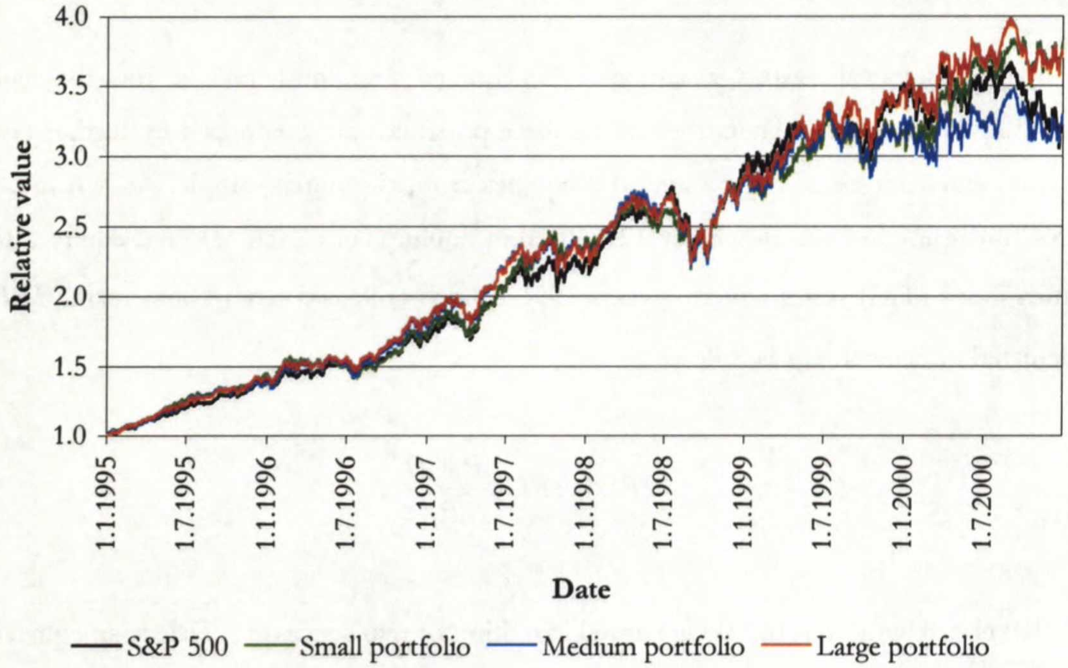


Figure 1 shows that the value development is rather uniform among the S&P 500 index and the portfolios until to the beginning of 1999. Thereafter the values start to diverge from each other more effectively.[21] At the end of the test period, the values of the small and large portfolios are about 3.7 times as large as at the beginning of the test period. On the other hand, for the S&P 500 index and medium portfolio the corresponding ratio is only about 3.2. Thus, the values of the small and large portfolios are roughly 16% higher than the S&P 500 and medium portfolio values at the of the test period. Overall, the development of the three portfolio values is such similar with the S&P 500 index that it supports the idea of mapping position onto the market; the market seems to be an important risk factor underlying in the portfolios. In the next three subchapters, the properties of the study's small, medium and large portfolios are presented in more detail.

[21] This effect could be seen better if logarithmic price levels were used instead of relative values. Figure 1 shows however relative values, so that the interesting value development of the S&P 500 index and the portfolios can be more easily observed.

4.2.1. Small portfolio

A total of 30 different S&P 500 constituents are represented in the small portfolio. In Table 1 below, sector representation of the portfolio is shown. The portfolio doesn't include any equities from Transportation sector, all other seven eligible S&P 500 sectors being represented. As can be noticed from Table 1, the sector weights experience considerable changes in the six-year test period, implying that also the individual equity weights change very much. Relatively speaking, Technology sector gains most in value and Basic Materials least in the small portfolio. Moreover, the sector weight rankings undergo changes. For instance, Utilities is the largest sector in the small portfolio at the beginning of the test period but only the fifth largest sector at the end of the period.

Table 1.
Small portfolio's general sector information

Sector	Number of equities	Weight 1.1.1995	Weight 31.12.2000
Basic Materials	4	12.99 %	5.24 %
Capital Goods	4	14.00 %	11.04 %
Energy	3	11.91 %	8.38 %
Financials	6	15.38 %	22.76 %
Health Care	3	12.11 %	18.76 %
Technology	6	14.22 %	23.23 %
Utilities	4	19.38 %	10.59 %
Total	30	100.00 %	100.00 %

Table 2 below reports the small portfolio's central characteristics of both daily as well as weekly returns during the test period.[22]

Table 2.
Small portfolio's return characteristics on test period 1995-2000

	Daily returns	Weekly returns
Sample mean	0.09 %	0.44 %
Sample standard deviation	1.04 %	2.17 %
Sample skewness	-0.31	-0.27
Sample kurtosis	7.06	3.96
Test statistic t for zero mean return	**3.27****	**7.78****
Test statistic A^* for normality of returns	**7.36****	**2.53****

** Significant at the 1% level.

[22] Sample skewness is calculated as $\frac{\sum_{i=1}^{s}(R_i-\bar{R})^3}{s\sigma^3}$ and sample kurtosis as $\frac{\sum_{i=1}^{s}(R_i-\bar{R})^4}{s\sigma^4}$.

The sample mean daily return of 0.09% is rather low, but still the null hypothesis of zero mean return is rejected at the significance level of 1%. When the weekly returns are concerned, the rejection of the null is even clearer, as could be expected. Thus, when backtesting the relative VaR measures of the small portfolio, the mean returns have to be taken into account.

For both daily and weekly portfolio returns, the sample skewness is slightly negative indicating that the return distribution isn't symmetric contrary to the properties of the normal distribution, and has more return observations on the right side of the mean than on the left side. Therefore, the distribution has relatively large returns below the mean, which isn't a desired feature when the variance-covariance approach is applied for VaR measures. Furthermore, the sample kurtosis is over three for the both returns and thus exceeds the kurtosis of the normal distribution. Consequently, the observed return distributions have thicker tails than the normal distribution should have, which is also an undesired feature and may well result in too low VaR measures in the study. The observed return distributions' skewness and kurtosis deviations from the ideal normal distribution values can be noticed from Appendix B, in which the observed distributions are shown together with the respective normal distributions. (Campbell, Lo and MacKinlay, 1997)

As the skewness and kurtosis numbers already indicate, the Anderson-Darling test rejects the null hypothesis of normally distributed portfolio returns. The rejection is made to both daily and weekly returns at the 1% significance level. The modified Anderson-Darling test statistics A^*:s imply that the small portfolio's weekly returns are closer to normality than the daily returns. This implication is further strengthened when Figure B1 is compared to Figure B2 in Appendix B.

4.2.2. Medium portfolio

The study's medium portfolio consists of 50 equity constituents of the S&P 500 index. Including two equities from Transportation sector, all the eight eligible S&P 500 sectors are represented contrary to the previously presented small portfolio. However, it should be noted that during the test period the weight of Transportation sector remains rather low, mainly under 5%, in the medium portfolio, and the inclusion of the sector doesn't result in considerable sector weight differences compared to the small portfolio. Nevertheless, the sector weights summarised in Table 3 below do differ quite much from the small portfolio's respective weights, except for the Basic Materials and Health Care sectors. Overall, the sector weights of the portfolio experience parallel changes as the small portfolio's weights in the test period. The only exception is Capital Goods, which has more weight at the end of the period than at the beginning; for the small portfolio, the situation is opposite.[23]

Table 3.
Medium portfolio's general sector information

Sector	Number of equities	Weight 1.1.1995	Weight 31.12.2000
Basic Materials	4	13.01 %	6.92 %
Capital Goods	8	14.99 %	19.58 %
Energy	5	7.99 %	4.19 %
Financials	11	20.97 %	22.83 %
Health Care	4	11.48 %	18.49 %
Technology	11	19.89 %	20.99 %
Transportation	2	4.73 %	2.51 %
Utilities	5	6.93 %	4.49 %
Total	50	100.00 %	100.00 %

The return characteristics of the medium portfolio shown below in Table 4 are similar to the small portfolio. The medium portfolio's sample means and standard deviations are very close to the respective small portfolio figures. Therefore, the *t*-test again rejects the null hypothesis of zero expected portfolio return for both daily and weekly returns at the 1% significance level.

Table 4.
Medium portfolio's return characteristics on test period 1995-2000

	Daily returns	Weekly returns
Sample mean	0.08 %	0.39 %
Sample standard deviation	1.08 %	2.39 %
Sample skewness	-0.26	-0.23
Sample kurtosis	5.92	4.16
Test statistic t for zero mean return	**2.78****	**6.32****
Test statistic A^* for normality of returns	**8.24****	**2.67****

** Significant at the 1% level.

[23] This is primarily due to the strong performances of the General Electric's and United Technologies' equities, which are included in the medium but not in the small portfolio.

The undesired features of negative skewness and excess kurtosis over the normal distribution exist also in the observed medium portfolio distribution. These features, especially the excess kurtosis, can be discovered from Appendix C's figures C1 and C2. Therefore, it comes as no surprise that the value of the modified Anderson-Darling test statistic A^* implies rejection of the normal distribution hypothesis for the both returns at the significance level of 1%. Similarly to the small portfolio, the test statistics as well as figures B1 and B2 indicate that the departure from normality is stronger among the daily than the weekly medium portfolio returns.

4.3.3. Large portfolio

On the basis of number of equities included, the large portfolio with its 100 equities from the S&P 500 constituents is diversified over three times as well as the small portfolio and twice as well as the medium portfolio. The all eight eligible S&P 500 sectors are represented in the large portfolio, as can be seen from Table 5 below. The sector weights of the portfolio differ to some extent from the other two portfolios' respective weights. Nevertheless, the large portfolio experiences same kind of sector weight changes in the test period as the small and medium portfolios do; this time the weight of Capital Goods remains almost unchanged.

Table 5.
Large portfolio's general sector information

Sector	Number of equities	Weight 1.1.1995	Weight 31.12.2000
Basic Materials	11	10.59 %	4.43 %
Capital Goods	15	17.96 %	17.16 %
Energy	8	13.92 %	10.35 %
Financials	22	19.25 %	23.88 %
Health Care	7	10.21 %	15.33 %
Technology	15	12.57 %	18.70 %
Transportation	5	4.40 %	2.50 %
Utilities	17	11.10 %	7.65 %
Total	100	100.00 %	100.00 %

The large portfolio shares the return characteristics of the other two portfolios. The sample means and standard deviations reported in the next page's Table 6 are fairly identical with the respective figures of the smaller portfolios. Consequently, the t-test statistics are above the critical values of the 99% confidence level, and the null of zero expected return is rejected for the both daily as well as weekly returns.

Table 6.
Large portfolio's return characteristics on test period 1995-2000

	Daily returns	Weekly returns
Sample mean	0.09 %	0.43 %
Sample standard deviation	1.06 %	2.25 %
Sample skewness	-0.21	-0.24
Sample kurtosis	6.61	4.47
Test statistic t for zero mean return	**3.19****	**7.49****
Test statistic A^* for normality of returns	**9.06****	**3.21****

** Significant at the 1% level.

The sample skewness is slightly negative also for the large portfolio and the sample kurtosis is above the ideal normal distribution level of three. These departures from normality can be seen from Appendix D, which shows the observed return distributions during the test period 1995-2000 alongside the corresponding normal distributions. The study's observed distributions' test statistics A^*:s achieve their highest values in the large portfolio for the both daily and weekly returns, and the Anderson-Darling test strongly rejects the portfolio return normality at the 1% significance level. Furthermore, imitating the findings of the two smaller portfolios, the large portfolio's test statistics and the figures of Appendix D indicate that the normality assumption suits better for weekly than daily returns.

5. Results

Performance of Traditional (7), Beta (16), Diagonal beta (17), Sector-beta (23) and Diagonal sector-beta (24) model during the test period 1995-2000 is reported in the following three subchapters. Firstly, only the portfolio return standard deviations used in different VaR models during the test period are under review. In the second subchapter, the output of the standard deviations, the VaR measures, are considered and the VaR model accuracy results are presented. Finally, the VaR accuracy results and findings are summarised in the third subchapter.

5.1. Standard deviation comparisons

Since all the accuracy differences among the study's five VaR models result from differing portfolio return standard deviation calculus, comparison of the respective standard deviations gives important indications on these accuracy differences. Consequently, the standard deviations used in the models during the test period are compared next by portfolios. The comparisons are made for the daily portfolio standard deviation values, but the general findings are identical also for the weekly standard deviations, as they are merely multiples of the daily standard deviations.

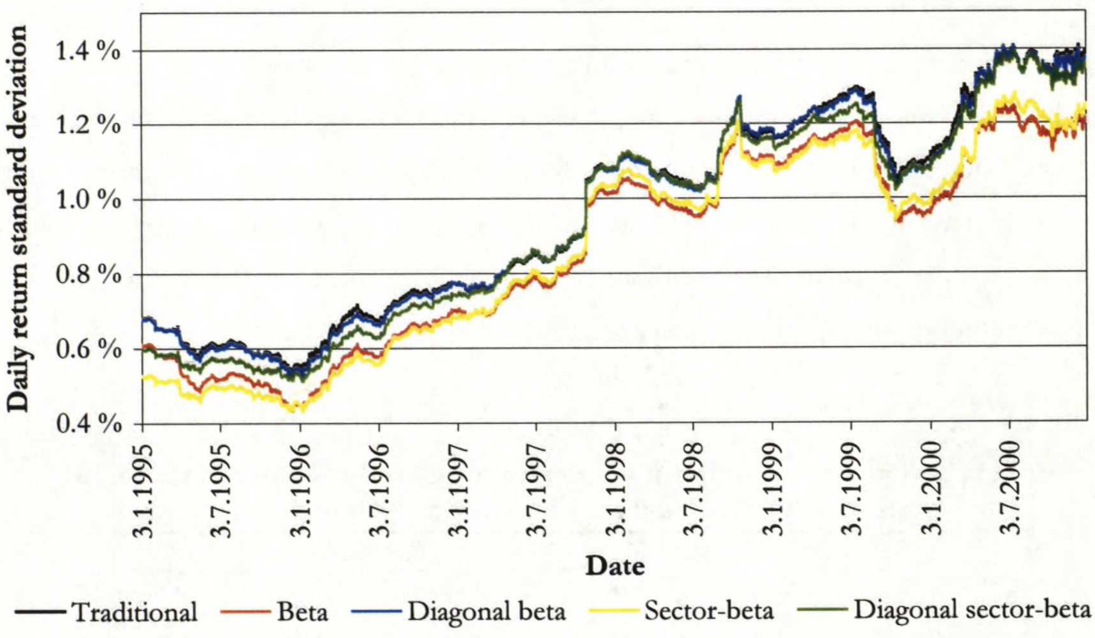
It's worthwhile to utilise graphical analysis for the standard deviations in order to find out evidence on the differences among the VaR models, and especially to see how these differences change during the test period. Thus, figures are presented that show the daily portfolio return standard deviations used in the five VaR models during the test period.

In addition to the figures, it's reported how well a standard deviation used in a beta mapping-based VaR model corresponds, on average, to a benchmark standard deviation that is calculated for Traditional model through Formula (4). These correspondences, or rather mean ratios, are obtained through the following two steps for each beta mapping-based VaR model and estimation period length: Firstly, the ratio of beta mapping-based standard deviation to corresponding traditionally calculated standard deviation is computed each day on the test period; the both standard deviations use identical estimation periods. Then, these ratios are averaged to obtain the mean ratio for the particular series.

5.1.1. Small portfolio

In Figure 2 below, it's shown how the standard deviations used in the VaR models develop when the small portfolio is concerned and the parameters of the models are estimated using 250 days mowing window. When the longer estimation period lengths of 500 and 1000 days are applied, development of the standard deviations gets smoother as significance of individual return observation diminishes and the ghosting feature loses its power; this phenomenon can be seen by comparing Figure 2 to Appendix E's figures E1 and E2. Nevertheless, on the whole the differences among the VaR models are rather similar in all the three figures. Therefore, the general findings from Figure 1 are valid also for the outcomes concerning the 500 and 1000 days estimation period lengths.

Figure 2.
Daily return standard deviations used in different VaR models during test period (small portfolio, estimation period length 250 days)



It can be noticed from Figure 2 that in each model the small portfolio's standard deviation, and thus also the respective VaR measures, experience a rough upward trend during the test period. The absolute standard deviation gaps between the VaR models vary to some extent during the period, implying that the accuracy differences, especially between the two mapping approaches, may be sensitive to the chosen backtesting period 1995-2000.

At the beginning of the test period in Figure 2, the standard deviations of the VaR models applying mapping onto the sector indices are lower than the corresponding standard deviations that result from mapping onto the market index. This is due to the fact that although the OLS beta regressions' R^2 statistics, which are not reported in the study, are throughout the test period higher for an sector index than for the market index regressions, the correlations between the different sector indices are too low to favour the VaR models that apply mapping onto the sector indices, as explained in Appendix A.[24] As time passes by in the test period, the correlation conditions change and advance the standard deviation performance of mapping onto the sector indices.

Furthermore, worth noticing from Figure 2 is Diagonal beta model's ability to use very identical standard deviations with Traditional model throughout the test period 1995-2000. Also, the standard deviations of Diagonal sector-beta model are in general fairly close to the benchmark values of Traditional model during the test period.

The mean ratios presented below in Table 7 confirm that the standard deviations used in the both diagonal models are on average close to Traditional model's corresponding values; Diagonal beta model is generally lacking only about half percent in the benchmark standard deviation. This implies that the residual term assumptions (9) are appropriate and that the VaR accuracy performances of Traditional, Diagonal beta and Diagonal sector-beta model should be rather similar for the small portfolio.

Table 7.
Test period's mean ratios of beta mapping-based standard deviation to traditionally calculated standard deviation (small portfolio)

	VaR model			
Estimation period length	Beta	Diagonal beta	Sector-beta	Diagonal sector-beta
250 days	89.64 %	99.43 %	88.92 %	96.92 %
500 days	89.74 %	99.50 %	88.44 %	96.56 %
1000 days	89.30 %	99.64 %	87.49 %	96.13 %

[24] For the study's OLS regressions, the R^2 statistic, defined as the square of the regression's correlation coefficient ρ_{12}, describes how much does the independent variable, the general market index return or an sector index return, explain of the dependent variable's, individual equity return's, total variation in the estimation period.

As reported in Table 7, in contrast to the diagonal models, the two plain beta models' standard deviations deviate relatively much, usually over the percentage, from the benchmark values. Therefore, their VaR accuracy level can easily differ from Traditional model. Overall, the mean ratios don't seem to be dependent on the estimation period length, indicating that the numbers of the VaR measure exceptions incur for the different models similarly irrespective of the estimation period length under consideration.

When the plain beta VaR models are compared to their respective diagonal models, it can be noted from Table 7 that inclusion of the residual terms, which represent the firm-specific risks, increases more Diagonal beta model's than Diagonal sector-beta model's standard deviations. This follows from the previously mentioned R^2 statistic differences between the sector index and general market index regressions, which imply higher residual standard deviations for VaR models that apply mapping onto the market index than for the VaR models that map positions onto the sector indices.

5.1.2. Medium portfolio

Development of the medium portfolio's daily return standard deviation during the test period 1995-2000 is shown in Figure 3, on the next page, for each VaR model that applies the 250 days estimation period length.[25] The graphs in Figure 3 show similar development to the small portfolio's graphs with an exception that the medium portfolio's standard deviation rises a bit more rapidly. Again, in the early part of the test period 1995-2000, the VaR models applying position mapping onto the market index use noticeably higher standard deviations than the corresponding VaR models that apply mapping onto the sector indices, and the situation changes when moved towards the end of the test period.

Of special interest are the standard deviation gaps between Traditional model and the two plain beta models. As can be noticed by comparing Figure 3 to Figure 2, these gaps are clearly more narrow for the medium portfolio than for the small portfolio. This is according to expectations and results from the diminishing importance of the firm-specific risks as the portfolio size increases. Therefore, compared to the small portfolio, by the

[25] Appendix F presents the corresponding figures for the 500 and 1000 days estimation period lengths, which share the same general characteristics with Figure 3 although the graphs are slightly smoother.

medium portfolio the VaR accuracy of Beta model and Sector-beta model should be more similar to Traditional model.

Figure 3.
Daily return standard deviations used in different VaR models during test period (medium portfolio, estimation period length 250 days)

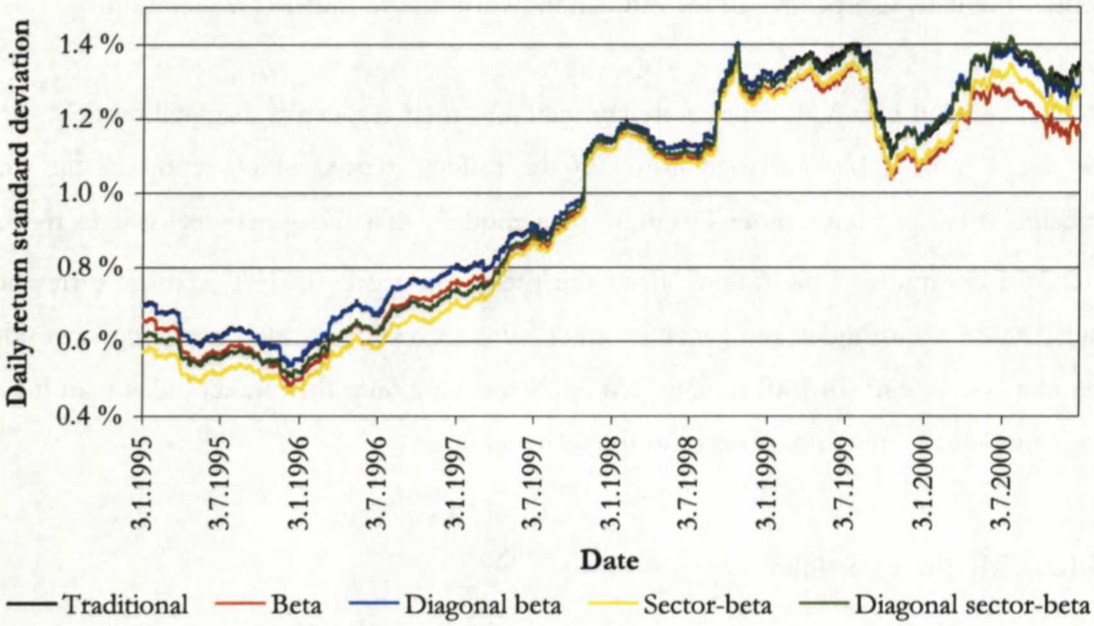


Table 8 below reports that during the test period 1995-2000 Beta model uses, on average, only about six percent lower standard deviation than Traditional model. On the other hand, Sector-beta model's mean ratios are a bit lower compared to Beta model and depend quite meaningfully on the estimation period length, the shortest estimation period length of 250 days providing the most competitive standard deviations.

Table 8.
Test period's mean ratios of beta mapping-based standard deviation to traditionally calculated standard deviation (medium portfolio)

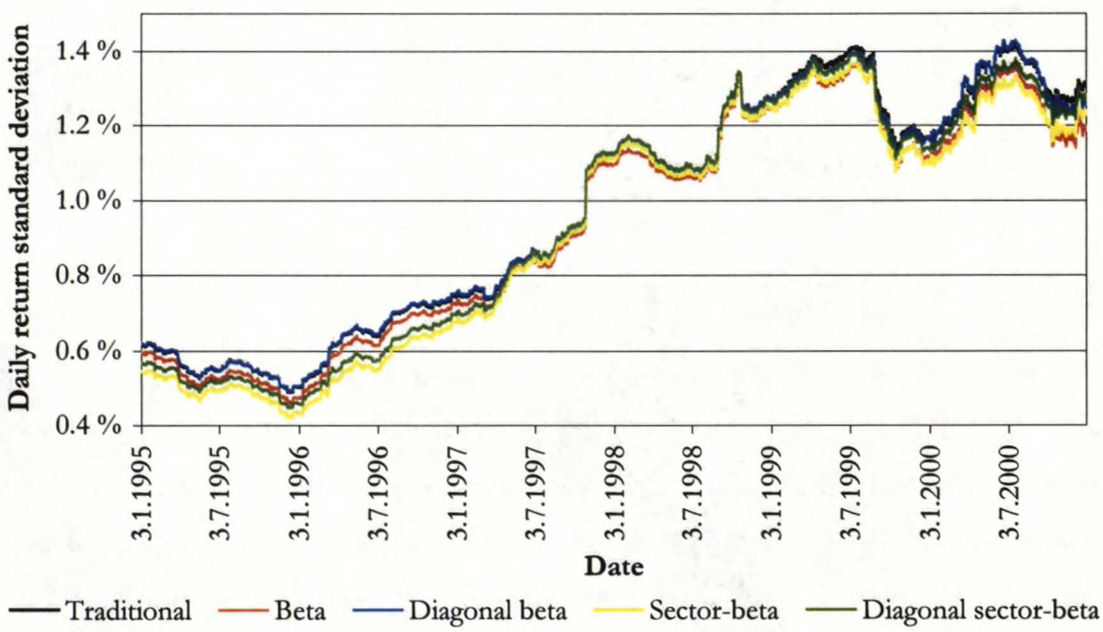
	VaR model			
Estimation period length	Beta	Diagonal beta	Sector-beta	Diagonal sector-beta
250 days	94.24 %	99.71 %	91.82 %	96.25 %
500 days	94.25 %	99.60 %	90.73 %	95.24 %
1000 days	93.77 %	99.34 %	88.92 %	93.71 %

Furthermore, the both diagonal models have fairly similar mean ratios for the medium as for the small portfolio: Diagonal beta model uses even more identical standard deviations with Traditional model, while Diagonal sector-beta model's correspondence decreases; this drop is mainly due Transportation sector's low correlation with the other seven sectors.

5.1.3. Large portfolio

By doubling the medium portfolio's number of equities, it's expected that the large portfolio's return standard deviations used in Beta and Sector-beta model further strengthen their correspondences to Traditional model's benchmark standard deviations. This is indeed what happens, as can be observed by comparing the standard deviation gaps in Figure 3 to the respective gaps in Figure 4 below, which presents the large portfolio's daily return standard deviations used during the test period in the five VaR models that estimate the necessary parameters applying 250 days moving window.[26]

Figure 4.
Daily return standard deviations used in different VaR models during test period (large portfolio, estimation period length 250 days)



---

[26] Similarly to the small and medium portfolio, the daily return standard deviation of the large portfolio develops somewhat smoother but otherwise quite identically with graphs of Figure 4, when 500 and 1000 days are applied in parameter estimation; Appendix G shows these standard deviation developments.

The general development of the large portfolio's standard deviation is very similar to the medium portfolio. The comparative standard deviation advancement phenomenon resulting from mapping equities onto the sector indices exists during the test period also in the large portfolio. In other words, Sector-beta and Diagonal sector-beta models use lower standard deviations than the respective VaR models applying position mapping onto the market index at the beginning of the period, but strengthen their comparative performance when moved longer in the test period. Overall, the standard deviation gaps are very narrow in Figure 4, and especially Diagonal beta model seems to use rather perfect standard deviation substitutes for Traditional model.

On the basis of the mean standard deviation ratios reported below in Table 9, Diagonal beta model is expected to have a very identical VaR accuracy performance with Traditional model for the study's large portfolio. Furthermore, Beta model may well turn out to be fairly equally accurate as Traditional model when the large portfolio is concerned, since its standard deviation misses in general only by four percent the corresponding benchmark value of Traditional model.

Table 9.
Test period's mean ratios of beta mapping-based standard deviation to traditionally calculated standard deviation (large portfolio)

	VaR model			
Estimation period length	Beta	Diagonal beta	Sector-beta	Diagonal sector-beta
250 days	96.43 %	100.01 %	93.63 %	96.36 %
500 days	96.30 %	99.95 %	93.47 %	96.29 %
1000 days	95.79 %	99.79 %	92.68 %	95.77 %

It's interesting to note from Table 9 that irrespective of the estimation period length, the both VaR models applying position mapping onto the market index use on average standard deviations that more effectively correspond to the benchmark values of Traditional model than either of the VaR models that map positions onto the sector indices. Thus, for the study's large portfolio, mapping of individual equities onto the market index provides presumably comprehensively superior VaR measures to the measures derived from mapping positions onto the sector indices.

5.2. VaR accuracy results

The study's VaR accuracy results for the five models are obtained by comparing the calculated 95% and 99% VaR measures to the observed daily and weekly portfolio return outcomes. The test period covers years 1995-2000 implying 1515 daily return observations and a total of 1511 weekly return observations consisting of five different weekly series.[27] During the test period, all the necessary parameters are estimated for each VaR model applying an estimation period of past 250, 500 and 1000 days.

As reported in Chapter 4.2, the t-test rejects the null of zero expected return at the 1% significance level for each portfolio's daily and weekly returns. Thus, it's appropriate to subtract in each portfolio return series a test period sample mean from an observed return outcome, and increase the absolute value of the portfolio loss which is compared to the respective relative VaR measure. As noted previously, the term exception is used when a mean-corrected portfolio loss exceeds the respective VaR measure during the test period.

A number of exceptions, the respective failure rate and a mean VaR overdraft are presented for each examined VaR measure series. In order to ensure that the results of the study can be easily interpreted irrespective of an amount invested into a portfolio, the accuracy testing is made by using solely the portfolio returns and thus omitting the initial portfolio values W :s in the VaR models. Consequently, the mean VaR overdraft, which gives economically important information on the level of the exceptions, is denoted in percentage units and calculated in the following way: Whenever an exception occurs during the test period, its absolute difference from the respective VaR measure is computed. The mean VaR overdraft is then achieved for the particular series by averaging the observed differences.

Statistical tests are applied to the observed VaR measure exceptions, as described in Chapter 3.7.3. Thus, the test statistics LR_{uc} for the null of correct unconditional coverage,

27 Five separate weekly return series have to be used to test properly the independence of weekly VaR measure exceptions, because using one sole weekly return series that has new weekly return each day, and not at five-day intervals, would lead to biased independence results. However, since it's assumed that the weekly returns and exceptions are independent, each weekly return series' test statistic LR_{ind} can be added, and only these added values are reported and tested in the study.

LR_{ind} for the null hypothesis of independent exceptions and LR_{cc} for the null of correct conditional coverage are reported. Whenever a null is rejected, the particular test statistic is bolded. Furthermore, asterisks are used to indicate the statistical significance of the rejection: One asterisk refers to the rejection at the 5% significance level and two asterisks at the 1% level. Next, the accuracy results are presented by the five VaR models, in their respective subchapters. Since accuracy results are partly, for example concerning mean VaR overdrafts, very similar for study's each VaR model, these kind of findings are commented primarily in the first subchapter that presents the accuracy results of Traditional model.

5.2.1. Traditional model

Accuracy results for Traditional model's daily VaR measures are presented in Table 10, on the next page. Thus, based on the test period results, accuracy of Traditional model's daily VaR measure seems to fairly dependent on the estimation period length and the VaR confidence level. The shortest estimation period length, 250 days, works most accurately during the test period in Traditional model's each VaR measure series. This indicates that application of the longer estimation period lengths of 500 and 1000 days leads to standard deviation estimates which react too slowly to the changing market conditions.

Traditional model is able to achieve the correct unconditional coverage for the most of the 95% VaR measure series, but when the 99% VaR measures are concerned, the model lacks badly in accuracy and the null hypotheses of correct unconditional coverage are rejected at the 1% significance level; in other words, the 99% VaR measures produced by Traditional model are systematically too low. Since it's reported in the previous chapter that the other four models use in general lower standard deviations than Traditional model uses, these results imply weak accuracy results also for the study's other VaR models' 99% VaR measures.

There are also accuracy differences between the three portfolios, as can be noticed from Table 10. However, these differences are rather minor and result probably by a chance. Mean VaR overdrafts seem to be quite portfolio-centric and higher for the 99% than for the 95% VaR measures.

Table 10.
Daily VaR accuracy results for Traditional model

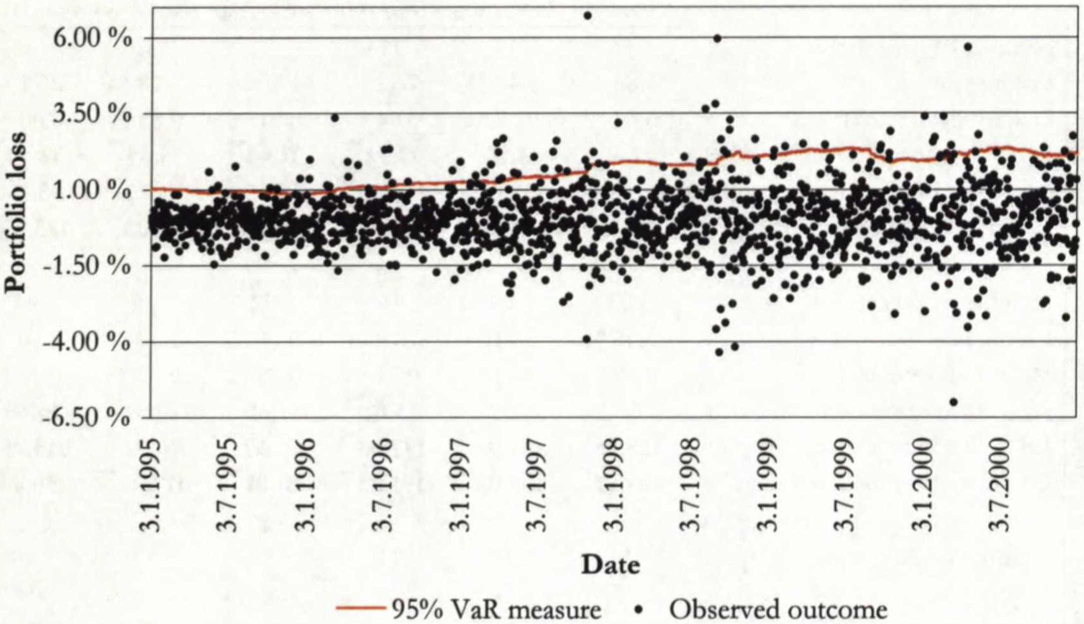
		VaR confidence level/estimation period length in days					
		95%/250	95%/500	95%/1000	99%/250	99%/500	99%/1000
Small portfolio	Number of exceptions	86	97	114	30	33	42
	Failure rate	5.68 %	6.40 %	7.52 %	1.98 %	2.18 %	2.77 %
	Mean VaR overdraft	0.64 %	0.62 %	0.64 %	0.81 %	0.83 %	0.80 %
	LR_{uc} for unconditional coverage	1.40	$\mathbf{5.79}^{*}$	$\mathbf{17.72}^{**}$	$\mathbf{11.44}^{**}$	$\mathbf{15.89}^{**}$	$\mathbf{32.44}^{**}$
	LR_{ind} for independent exceptions	$\mathbf{126.32}^{**}$	$\mathbf{138.57}^{**}$	$\mathbf{162.88}^{**}$	$\mathbf{68.70}^{**}$	$\mathbf{74.13}^{**}$	$\mathbf{93.15}^{**}$
	LR_{cc} for conditional coverage	$\mathbf{127.72}^{**}$	$\mathbf{144.36}^{**}$	$\mathbf{180.61}^{**}$	$\mathbf{80.14}^{**}$	$\mathbf{90.03}^{**}$	$\mathbf{125.59}^{**}$
Medium portfolio	Number of exceptions	83	90	109	32	36	44
	Failure rate	5.48 %	5.94 %	7.19 %	2.11 %	2.38 %	2.90 %
	Mean VaR overdraft	0.69 %	0.69 %	0.71 %	0.78 %	0.79 %	0.78 %
	LR_{uc} for unconditional coverage	0.71	2.67	$\mathbf{13.61}^{**}$	$\mathbf{14.35}^{**}$	$\mathbf{20.91}^{**}$	$\mathbf{36.68}^{**}$
	LR_{ind} for independent exceptions	$\mathbf{125.86}^{**}$	$\mathbf{135.41}^{**}$	$\mathbf{177.63}^{**}$	$\mathbf{66.67}^{**}$	$\mathbf{96.92}^{**}$	$\mathbf{113.75}^{**}$
	LR_{cc} for conditional coverage	$\mathbf{126.57}^{**}$	$\mathbf{138.08}^{**}$	$\mathbf{191.23}^{**}$	$\mathbf{81.01}^{**}$	$\mathbf{117.82}^{**}$	$\mathbf{150.43}^{**}$
Large portfolio	Number of exceptions	81	87	110	28	33	46
	Failure rate	5.35 %	5.74 %	7.26 %	1.85 %	2.18 %	3.04 %
	Mean VaR overdraft	0.67 %	0.68 %	0.68 %	0.89 %	0.83 %	0.74 %
	LR_{uc} for unconditional coverage	0.37	1.68	$\mathbf{14.39}^{**}$	$\mathbf{8.81}^{**}$	$\mathbf{15.89}^{**}$	$\mathbf{41.12}^{**}$
	LR_{ind} for independent exceptions	$\mathbf{127.83}^{**}$	$\mathbf{130.57}^{**}$	$\mathbf{185.29}^{**}$	$\mathbf{55.78}^{**}$	$\mathbf{89.14}^{**}$	$\mathbf{125.00}^{**}$
	LR_{cc} for conditional coverage	$\mathbf{128.20}^{**}$	$\mathbf{132.25}^{**}$	$\mathbf{199.68}^{**}$	$\mathbf{64.59}^{**}$	$\mathbf{105.03}^{**}$	$\mathbf{166.12}^{**}$

*Significant at the 5% level.
**Significant at the 1% level.

Although Table 10 above positively reports that Traditional model produces in the test period 1995-2000 daily 95% VaR measures that are on average correct and are not exceeded statistically significantly more or less than 5% of times, it unfortunately also reports that the correct conditional coverage isn't achieved for any of the 95% or 99% VaR measure series, since exceptions don't occur independently enough. This is a serious drawback and decreases attractiveness of Traditional model for daily VaR measures.

Next page's Figure 5 shows an example of how the exceptions are more clustered than evenly distributed during the test period for the large portfolio's daily 95% VaR measures applying the 250 days estimation period length; the portfolio outcomes in the figure are mean-corrected. As can be concluded from the figure, the standard deviation underlying in the VaR measure doesn't adapt effectively enough to the changing market circumstances.

Figure 5.
Daily 95% VaR measures and observed outcomes during test period 1995-2000 (Traditional model, estimation period length 250 days, large portfolio)



Traditional model's VaR measure accuracy results for the weekly portfolio returns, which are presented on the next page in Table 11, show similar features as for the daily portfolio returns. In other words, none of the 99% VaR measures provides statistically correct unconditional coverage and the shorter the estimation period the better the accuracy of the 95% VaR measure.

Contrary to the daily VaR measure examination, the mean VaR overdrafts for the weekly VaR measures are in general fairly equal between the 95% and 99% VaR confidence levels. Furthermore, the mean weekly VaR overdrafts are a bit higher and show more variety than the corresponding mean daily VaR overdrafts.

During the test period 1995-2000, the most significant accuracy difference between the daily and weekly VaR measures of Traditional model is the independence of the exceptions. As previously pointed out, the exceptions cluster quite strongly for the daily VaR measures, but this phenomenon is weaker for the weekly VaR measures.

Table 11.
Weekly VaR accuracy results for Traditional model

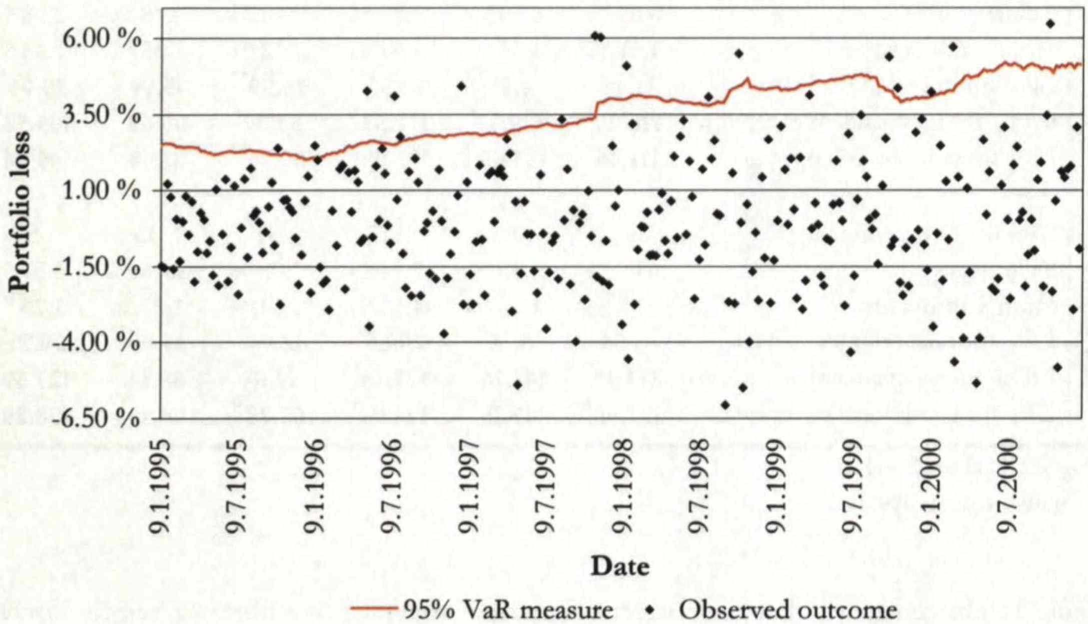
		VaR confidence level/estimation period length in days					
		95%/250	95%/500	95%/1000	99%/250	99%/500	99%/1000
Small portfolio	Number of exceptions	88	92	112	27	33	43
	Failure rate	5.82 %	6.09 %	7.41 %	1.79 %	2.18 %	2.85 %
	Mean VaR overdraft	1.15 %	1.24 %	1.29 %	1.25 %	1.22 %	1.28 %
	LR_{uc} for unconditional coverage	2.06	3.54	**16.22****	**7.66****	**15.99****	**34.69****
	LR_{ind} for independent exceptions	**117.81***	**129.14****	**169.35****	**50.27****	**58.52****	**99.49****
	LR_{cc} for conditional coverage	**119.86***	**132.67****	**185.57****	**57.93****	**74.51****	**134.18****
Medium portfolio	Number of exceptions	86	95	119	33	33	45
	Failure rate	5.69 %	6.29 %	7.88 %	2.18 %	2.18 %	2.98 %
	Mean VaR overdraft	1.31 %	1.33 %	1.34 %	1.12 %	1.35 %	1.43 %
	LR_{uc} for unconditional coverage	1.46	**4.89***	**22.56****	**15.99****	**15.99****	**39.04****
	LR_{ind} for independent exceptions	**110.00***	**129.77***	**172.53****	**53.37***	**65.69****	**105.52****
	LR_{cc} for conditional coverage	**111.46***	**134.66****	**195.09****	**69.36****	**81.68****	**144.56****
Large portfolio	Number of exceptions	92	96	117	31	33	50
	Failure rate	6.09 %	6.35 %	7.74 %	2.05 %	2.18 %	3.31 %
	Mean VaR overdraft	1.16 %	1.25 %	1.31 %	1.14 %	1.27 %	1.24 %
	LR_{uc} for unconditional coverage	3.54	**5.39***	**20.66****	**12.94****	**15.99****	**50.71****
	LR_{ind} for independent exceptions	**121.41***	**141.76****	**177.38****	**49.84***	**58.14****	**127.59****
	LR_{cc} for conditional coverage	**124.95***	**147.15****	**198.04****	**62.78****	**74.13****	**178.29****

* Significant at the 5% level.
** Significant at the 1% level.

Table 11 above reports that the correct conditional coverage hypothesis is rejected only at the 5% level of significance for the 95% weekly VaR measures applying 250 days estimation period length. Also, a couple of other VaR measure series in Table 11 have their exception independence test statistics LR_{ind} :s below the critical values corresponding to the 99% confidence level, although these series fail to be otherwise statistically significantly accurate.

None of Traditional model's VaR measure series is able to produce in the study correct conditional coverage that isn't rejected at least at the 5% significance level. Thus, keeping in mind the examination of Kupiec (1995) that indicates high type 2 error levels for the likelihood ratio tests applying the 99% confidence level, the exceptions may well in reality be dependent also for the weekly VaR measures of Traditional model. Some confirmation for this possibility is achieved, when the exceptions in the next page's Figure 6 are examined. Although the null hypothesis of independent exceptions isn't rejected at the 1% significance level for the small portfolio's weekly VaR measures applying the 250 days estimation period length, Figure 6 shows that the exceptions are not so evenly distributed during the test period; only one weekly series is presented in the figure, but the other four series' exceptions have similar characteristics.

Figure 6.
One weekly series' 95% VaR measures and observed outcomes during test period (Traditional model, estimation period length 250 days, small portfolio)



5.2.2. Beta model

Since it's reported earlier in the study that the standard deviations used in Beta models are in general below the corresponding values of Traditional model that has failure rates above the ideal levels of 5% and 1% during the test period, the even higher failure rates of Beta model aren't any surprise. Furthermore, as could be expected, the shortest estimation period of 250 days is again the most effective one and the 99% VaR measures produced by Beta model are on the whole really poor. Also, the mean VaR overdrafts of Beta model share the general characteristics of Traditional model. Beta model's accuracy results for the daily VaR measures are presented on the next page in Table 12.

Table 12.
Daily VaR accuracy results for Beta model

		VaR confidence level/estimation period length in days					
		95%/250	95%/500	95%/1000	99%/250	99%/500	99%/1000
Small portfolio	Number of exceptions	119	121	140	41	43	60
	Failure rate	7.85 %	7.99 %	9.24 %	2.71 %	2.84 %	3.96 %
	Mean VaR overdraft	0.60 %	0.64 %	0.66 %	0.78 %	0.81 %	0.72 %
	LR_{uc} for unconditional coverage	22.31**	24.28**	46.39**	30.38**	34.54**	76.82**
	LR_{ind} for independent exceptions	201.89**	193.77**	233.90**	107.77**	116.00**	166.27**
	LR_{cc} for conditional coverage	224.21**	218.05**	280.29**	138.15**	150.53**	243.08**
Medium portfolio	Number of exceptions	93	106	126	38	40	53
	Failure rate	6.14 %	7.00 %	8.32 %	2.51 %	2.64 %	3.50 %
	Mean VaR overdraft	0.70 %	0.66 %	0.69 %	0.76 %	0.81 %	0.75 %
	LR_{uc} for unconditional coverage	3.87*	11.37**	29.50**	24.54**	28.38**	58.01**
	LR_{ind} for independent exceptions	139.95**	157.65**	215.72**	91.95**	107.31**	138.49**
	LR_{cc} for conditional coverage	143.82**	169.02**	245.23**	116.49**	135.69**	196.49**
Large portfolio	Number of exceptions	89	94	116	32	36	53
	Failure rate	5.87 %	6.20 %	7.66 %	2.11 %	2.38 %	3.50 %
	Mean VaR overdraft	0.66 %	0.68 %	0.70 %	0.84 %	0.82 %	0.71 %
	LR_{uc} for unconditional coverage	2.32	4.31*	19.50**	14.35**	20.91**	58.01**
	LR_{ind} for independent exceptions	140.69**	145.99**	200.46**	82.19**	99.14**	145.13**
	LR_{cc} for conditional coverage	143.01**	150.30**	219.96**	96.54**	120.05**	203.13**

*Significant at the 5% level.

**Significant at the 1% level.

It can be observed from Table 12 that the daily VaR measure accuracy of Beta model is during the test period rather dependent on the portfolio size. This results from the portfolios' differing abilities to diversify the firm-specific risk, as noticed previously in Chapter 5.1. Especially the accuracy difference between the small portfolio and the other two portfolios is quite meaningful. More precisely, likelihood ratio tests reject the correct coverage and exception independence hypotheses for each small portfolio's VaR measure series at the 1% significance level. On the other hand, based on the likelihood ratio tests for correct unconditional coverage, the daily 95% VaR measures are accurate for the medium and large portfolio, when the shortest estimation period length of 250 days is applied for the parameters; for the medium portfolio, the correct unconditional coverage hypothesis is though rejected at the significance level of 5%. However, similarly to Traditional model, none of Beta model's daily VaR measure series leads to acceptation of correct conditional coverage hypothesis, because the exceptions don't occur statistically independently during the test period 1995-2000.

Table 13 below presents the accuracy results for Beta model's weekly VaR measures. As can be seen from the table, under the weekly VaR measure examination, also the medium portfolio's correct unconditional coverage hypotheses are rejected at the 1% significance level for each VaR measure series in addition to the small portfolio's respective hypotheses. Furthermore, none of the series avoids the rejection in any of the three likelihood ratio tests at the significance level of 5%.

Table 13.
Weekly VaR accuracy results for Beta model

		VaR confidence level/estimation period length in days					
		95%/250	95%/500	95%/1000	99%/250	99%/500	99%/1000
Small portfolio	Number of exceptions	113	123	143	44	53	71
	Failure rate	7.48 %	8.14 %	9.46 %	2.91 %	3.51 %	4.70 %
	Mean VaR overdraft	1.18 %	1.20 %	1.29 %	1.15 %	1.13 %	1.12 %
	LR_{uc} for unconditional coverage	**17.07****	**26.58****	**50.80****	**36.84****	**58.21****	**110.05****
	LR_{ind} for independent exceptions	**167.86****	**187.39****	**240.90****	**106.98****	**154.73****	**230.50****
	LR_{cc} for conditional coverage	**184.93****	**213.98****	**291.70****	**143.82****	**212.94****	**340.55****
Medium portfolio	Number of exceptions	104	113	143	36	42	57
	Failure rate	6.88 %	7.48 %	9.46 %	2.38 %	2.78 %	3.77 %
	Mean VaR overdraft	1.25 %	1.28 %	1.28 %	1.27 %	1.27 %	1.34 %
	LR_{uc} for unconditional coverage	**10.14****	**17.07****	**50.80****	**21.02****	**32.58****	**68.76****
	LR_{ind} for independent exceptions	**132.68***	**162.80****	**221.53****	**67.83****	**95.94****	**171.95****
	LR_{cc} for conditional coverage	**142.82****	**179.87****	**272.33****	**88.85****	**128.52****	**240.71****
Large portfolio	Number of exceptions	97	106	128	36	41	54
	Failure rate	6.42 %	7.02 %	8.47 %	2.38 %	2.71 %	3.57 %
	Mean VaR overdraft	1.21 %	1.23 %	1.31 %	1.11 %	1.15 %	1.29 %
	LR_{uc} for unconditional coverage	**5.91***	**11.54****	**32.01****	**21.02****	**30.52****	**60.79****
	LR_{ind} for independent exceptions	**122.59***	**159.15****	**204.47****	**66.38****	**95.35****	**147.83****
	LR_{cc} for conditional coverage	**128.50***	**170.69****	**236.48****	**87.40****	**125.88****	**208.62****

* Significant at the 5% level.
** Significant at the 1% level.

It's noticeable from Table 13 that Beta model's 95% VaR measure exceptions are acceptably independent under the confidence level of 99% only for the medium and large portfolio when the 250 days estimation period is applied. Moreover, the correct conditional coverage hypothesis isn't rejected at the 1% significance level only for the large portfolio's weekly 95% VaR measures applying the shortest estimation period length.

5.2.3. Diagonal beta model

As the standard deviation comparisons presented in Chapter 5.1 already indicate, the inclusion of the firm-specific risks, in the form of the OLS regressions' residual term standard deviations $\sigma_{\varepsilon,i}$:s, improves the VaR measure accuracy when individual equity positions are mapped onto the general market index, particularly for the small portfolio. The 99% VaR measures remain to be clearly inaccurate also for Diagonal beta model, but especially the 95% VaR measures applying the shortest estimation period length of 250 days are fairly accurate during the test period 1995-2000. Table 14 below reports the test period's accuracy results for Diagonal beta model's daily VaR measures.

Table 14.
Daily VaR accuracy results for Diagonal beta model

		VaR confidence level/estimation period length in days					
		95%/250	95%/500	95%/1000	99%/250	99%/500	99%/1000
Small portfolio	Number of exceptions	87	98	116	30	32	42
	Failure rate	5.74 %	6.47 %	7.66 %	1.98 %	2.11 %	2.77 %
	Mean VaR overdraft	0.64 %	0.62 %	0.64 %	0.83 %	0.87 %	0.81 %
	LR_{uc} for unconditional coverage	1.68	**6.32***	**19.50****	**11.44****	**14.35****	**32.44****
	LR_{ind} for independent exceptions	**125.59****	**144.74****	**167.06****	**68.70****	**70.06****	**93.15****
	LR_{cc} for conditional coverage	**127.27****	**151.07****	**186.57****	**80.14****	**84.41****	**125.59****
Medium portfolio	Number of exceptions	82	91	110	32	36	45
	Failure rate	5.41 %	6.01 %	7.26 %	2.11 %	2.38 %	2.97 %
	Mean VaR overdraft	0.70 %	0.69 %	0.71 %	0.78 %	0.79 %	0.77 %
	LR_{uc} for unconditional coverage	0.53	3.04	**14.39****	**14.35****	**20.91****	**38.88****
	LR_{ind} for independent exceptions	**122.01****	**133.71****	**175.20****	**66.67****	**96.92****	**116.53****
	LR_{cc} for conditional coverage	**122.53****	**136.75****	**189.59****	**81.01****	**117.82****	**155.40****
Large portfolio	Number of exceptions	81	87	112	28	33	46
	Failure rate	5.35 %	5.74 %	7.39 %	1.85 %	2.18 %	3.04 %
	Mean VaR overdraft	0.67 %	0.68 %	0.67 %	0.89 %	0.82 %	0.74 %
	LR_{uc} for unconditional coverage	0.37	1.68	**16.02****	**8.81****	**15.89****	**41.12****
	LR_{ind} for independent exceptions	**127.83****	**130.57****	**191.45****	**55.78****	**89.14****	**123.22****
	LR_{cc} for conditional coverage	**128.20****	**132.25****	**207.47****	**64.59****	**105.03****	**164.34****

*Significant at the 5% level.

**Significant at the 1% level.

It's reported in Table 14 that Diagonal beta model's daily 95% VaR measures applying the shortest estimation period length provide statistically correct unconditional coverage for each portfolio; also, when 500 days is applied in the parameter estimation, the correct unconditional coverage hypothesis is rejected at the 5% significance level only for the small portfolio's daily 95% VaR measures. Unfortunately, the exceptions are so clustered during the test period that the correct conditional hypothesis isn't accepted for any of Diagonal beta model's daily VaR measure series.

Contrary to the daily VaR measures, Diagonal beta model is capable in producing weekly VaR measures which confront exceptions quite steadily during the test period. Thus, for many weekly VaR measure series of Diagonal beta model, the exception independence hypothesis is rejected only at the 5% significance level, as can be observed from Table 15 below.

Table 15.
Weekly VaR accuracy results for Diagonal beta model

		VaR confidence level/estimation period length in days					
		95%/250	95%/500	95%/1000	99%/250	99%/500	99%/1000
Small portfolio	Number of exceptions	91	93	112	28	33	44
	Failure rate	6.02 %	6.15 %	7.41 %	1.85 %	2.18 %	2.91 %
	Mean VaR overdraft	1.13 %	1.25 %	1.30 %	1.24 %	1.25 %	1.26 %
	LR_{uc} for unconditional coverage	3.13	**3.96***	**16.22****	**8.87****	**15.99****	**36.84****
	LR_{ind} for independent exceptions	**124.92***	**127.97****	**172.99****	**53.61****	**61.01****	**102.38****
	LR_{cc} for conditional coverage	**128.05****	**131.93****	**189.21****	**62.49****	**77.00****	**139.22****
Medium portfolio	Number of exceptions	86	97	122	33	33	46
	Failure rate	5.69 %	6.42 %	8.07 %	2.18 %	2.18 %	3.04 %
	Mean VaR overdraft	1.31 %	1.31 %	1.32 %	1.12 %	1.36 %	1.41 %
	LR_{uc} for unconditional coverage	1.46	**5.91***	**25.55****	**15.99****	**15.99****	**41.28****
	LR_{ind} for independent exceptions	**110.00***	**131.12***	**180.51****	**53.37***	**65.69****	**112.05****
	LR_{cc} for conditional coverage	**111.46***	**137.03****	**206.06****	**69.36****	**81.68****	**153.34****
Large portfolio	Number of exceptions	91	98	117	31	32	50
	Failure rate	6.02 %	6.49 %	7.74 %	2.05 %	2.12 %	3.31 %
	Mean VaR overdraft	1.17 %	1.22 %	1.31 %	1.13 %	1.30 %	1.24 %
	LR_{uc} for unconditional coverage	3.13	**6.45***	**20.66****	**12.94****	**14.44****	**50.71****
	LR_{ind} for independent exceptions	**118.07***	**138.30****	**177.20****	**46.82***	**54.59****	**127.59****
	LR_{cc} for conditional coverage	**121.20***	**144.75****	**197.86****	**59.77****	**69.02****	**178.29****

* Significant at the 5% level.
** Significant at the 1% level.

Although the failure rates for weekly VaR measures are in Table 15 slightly higher than the corresponding failure rates in Table 14 for the daily VaR measures, the correct unconditional coverage hypothesis isn't rejected for any of the weekly 95% measures applying the 250 days estimation period length. Consequently, the correct conditional coverage hypothesis is rejected only at the 5% significance level for the medium and large portfolio's weekly 95% VaR measures that apply the shortest estimation period length; for the small portfolio, the hypothesis is rejected narrowly also at the significance level of 1%.

5.2.4. Sector-beta model

The VaR measure accuracy results for Sector-beta model are in general very poor, which is not surprising by noticing the standard deviation comparisons in Chapter 5.1 and the earlier accuracy results of Traditional model. Nevertheless, Sector-beta model is able to produce such an accurate failure rate for the large portfolio's daily 95% VaR measures applying the 250 days parameter estimation length that the correct unconditional coverage hypothesis isn't rejected even at the 5% significance level. The otherwise rather unpleasant accuracy results for the daily VaR measures can be noted from Table 16 below.

Table 16.
Daily VaR accuracy results for Sector-beta model

		VaR confidence level/estimation period length in days					
		95%/250	95%/500	95%/1000	99%/250	99%/500	99%/1000
Small portfolio	Number of exceptions	117	128	146	42	48	64
	Failure rate	7.72 %	8.45 %	9.64 %	2.77 %	3.17 %	4.22 %
	Mean VaR overdraft	0.61 %	0.61 %	0.64 %	0.76 %	0.74 %	0.69 %
	LR_{uc} for unconditional coverage	**20.42****	**31.72****	**54.59****	**32.44****	**45.73****	**88.34****
	LR_{ind} for independent exceptions	**192.62****	**208.14****	**247.15****	**106.96****	**129.41****	**182.41****
	LR_{cc} for conditional coverage	**213.04****	**239.85****	**301.74****	**139.39****	**175.14****	**270.75****
Medium portfolio	Number of exceptions	100	115	138	39	44	61
	Failure rate	6.60 %	7.59 %	9.11 %	2.57 %	2.90 %	4.03 %
	Mean VaR overdraft	0.68 %	0.65 %	0.69 %	0.79 %	0.82 %	0.74 %
	LR_{uc} for unconditional coverage	**7.46****	**18.60****	**43.78****	**26.44****	**36.68****	**79.65****
	LR_{ind} for independent exceptions	**152.69****	**184.36****	**257.08****	**96.09****	**122.43****	**173.12****
	LR_{cc} for conditional coverage	**160.15****	**202.96****	**300.86****	**122.52****	**159.12****	**252.76****
Large portfolio	Number of exceptions	90	98	128	34	37	56
	Failure rate	5.94 %	6.47 %	8.45 %	2.24 %	2.44 %	3.70 %
	Mean VaR overdraft	0.69 %	0.68 %	0.66 %	0.84 %	0.86 %	0.72 %
	LR_{uc} for unconditional coverage	2.67	**6.32***	**31.72****	**17.51****	**22.70****	**65.85****
	LR_{ind} for independent exceptions	**141.98****	**153.08****	**222.31****	**87.63****	**101.94****	**159.32****
	LR_{cc} for conditional coverage	**144.65****	**159.41****	**254.02****	**105.13****	**124.64****	**225.16****

* Significant at the 5% level.

** Significant at the 1% level.

When the weekly VaR measures of Sector-beta model are concerned, the accuracy results shown below in Table 17 are even worse than for the respective daily VaR measures. The failure rates are far above the ideal levels of 5% and 1%. Therefore, the correct unconditional coverage hypothesis is rejected for each VaR measure series at the 1% significance level. In fact, there's only one hypothesis that isn't rejected at the 1% level of significance and that's the independence of exceptions hypothesis for the large portfolio's weekly 95% VaR measures applying the shortest estimation period length; this hypothesis is rejected only at the 5% significance level.

Table 17.
Weekly VaR accuracy results for Sector-beta model

		VaR confidence level/estimation period length in days					
		95%/250	95%/500	95%/1000	99%/250	99%/500	99%/1000
Small portfolio	Number of exceptions	114	124	148	46	52	70
	Failure rate	7.54 %	8.21 %	9.79 %	3.04 %	3.44 %	4.63 %
	Mean VaR overdraft	1.18 %	1.21 %	1.28 %	1.11 %	1.19 %	1.20 %
	LR_{uc} for unconditional coverage	**17.94****	**27.63****	**57.86****	**41.28****	**55.67****	**106.90****
	LR_{ind} for independent exceptions	**167.60****	**191.60****	**254.02****	**112.68****	**150.20****	**232.16****
	LR_{cc} for conditional coverage	**185.53****	**219.23****	**311.87****	**153.97****	**205.87****	**339.06****
Medium portfolio	Number of exceptions	110	123	157	44	54	69
	Failure rate	7.28 %	8.14 %	10.39 %	2.91 %	3.57 %	4.57 %
	Mean VaR overdraft	1.25 %	1.29 %	1.30 %	1.14 %	1.14 %	1.31 %
	LR_{uc} for unconditional coverage	**14.58****	**26.58****	**71.49****	**36.84****	**60.79****	**103.77****
	LR_{ind} for independent exceptions	**156.85****	**188.99****	**277.85****	**93.09****	**139.84****	**216.18****
	LR_{cc} for conditional coverage	**171.44****	**215.57****	**349.34****	**129.93****	**200.63****	**319.96****
Large portfolio	Number of exceptions	105	114	136	41	44	59
	Failure rate	6.95 %	7.54 %	9.00 %	2.71 %	2.91 %	3.90 %
	Mean VaR overdraft	1.19 %	1.23 %	1.32 %	1.11 %	1.20 %	1.31 %
	LR_{uc} for unconditional coverage	**10.83****	**17.94****	**41.58****	**30.52****	**36.84****	**74.26****
	LR_{ind} for independent exceptions	**141.06***	**167.51****	**223.11****	**87.78****	**111.35****	**168.97****
	LR_{cc} for conditional coverage	**151.89****	**185.45****	**264.69****	**118.30****	**148.19****	**243.23****

* Significant at the 5% level.
** Significant at the 1% level.

5.2.5. Diagonal sector-beta model

As expected, the consideration of the firm-specific risks in the standard deviation calculus increases the accuracy of mapping positions onto the sector indices. The 99% VaR measures remain inappropriate, but Diagonal sector-beta model is able to output quite accurate measures for the 95% VaR confidence level. The correct conditional coverage hypothesis isn't however accepted for any Diagonal sector-beta model's daily 95% VaR measures due to the clustered exceptions during the test period. Nevertheless, some daily 95% VaR measures are on average so accurate during the test period that the correct unconditional coverage hypothesis is accepted, as can be seen from Table 18 below. To be precise, application of the shortest estimation period length results in the rejection of the hypothesis only for the small portfolio's VaR measures at the 5% significance level. Also, Diagonal sector-beta model's daily 95% VaR measures that apply the 500 days estimation period length seem to provide correct unconditional coverage for the large portfolio's worst returns.

Table 18.
Daily VaR accuracy results for Diagonal sector-beta model

		VaR confidence level/estimation period length in days					
		95%/250	95%/500	95%/1000	99%/250	99%/500	99%/1000
Small portfolio	Number of exceptions	96	104	121	32	33	46
	Failure rate	6.34 %	6.86 %	7.99 %	2.11 %	2.18 %	3.04 %
	Mean VaR overdraft	0.61 %	0.62 %	0.64 %	0.81 %	0.89 %	0.79 %
	LR_{uc} for unconditional coverage	**5.27***	**9.98****	**24.28****	**14.35****	**15.89****	**41.12****
	LR_{ind} for independent exceptions	**134.72****	**150.74****	**181.68****	**81.46****	**75.49****	**113.61****
	LR_{cc} for conditional coverage	**139.99****	**160.72****	**205.96****	**95.80****	**91.38****	**154.73****
Medium portfolio	Number of exceptions	89	99	125	37	41	56
	Failure rate	5.87 %	6.53 %	8.25 %	2.44 %	2.71 %	3.70 %
	Mean VaR overdraft	0.69 %	0.69 %	0.69 %	0.74 %	0.79 %	0.72 %
	LR_{uc} for unconditional coverage	2.32	**6.88****	**28.42****	**22.70****	**30.38****	**65.85****
	LR_{ind} for independent exceptions	**130.17****	**151.82****	**213.47****	**90.05****	**108.86****	**147.28****
	LR_{cc} for conditional coverage	**132.49****	**158.70****	**241.89****	**112.75****	**139.24****	**213.13****
Large portfolio	Number of exceptions	87	91	116	32	35	51
	Failure rate	5.74 %	6.01 %	7.66 %	2.11 %	2.31 %	3.37 %
	Mean VaR overdraft	0.67 %	0.70 %	0.68 %	0.84 %	0.85 %	0.74 %
	LR_{uc} for unconditional coverage	1.68	3.04	**19.50****	**14.35****	**19.18****	**52.97****
	LR_{ind} for independent exceptions	**134.73****	**142.68****	**202.81****	**75.43****	**98.78****	**141.62****
	LR_{cc} for conditional coverage	**136.41****	**145.73****	**222.31****	**89.77****	**117.95****	**194.60****

* Significant at the 5% level.

** Significant at the 1% level.

Diagonal sector-beta model's accuracy results for the weekly VaR measures are presented on the next page. A bit surprisingly, concerning the weekly 95% VaR measures that apply the shortest estimation period length of 250 days, the correct unconditional coverage hypothesis is rejected at the 1% significance level merely for the large portfolio, while the hypothesis is rejected only at the significance level of 5% for the small and medium portfolio.

Table 19.
Weekly VaR accuracy results for Diagonal sector-beta model

		VaR confidence level/estimation period length in days					
		95%/250	95%/500	95%/1000	99%/250	99%/500	99%/1000
Small portfolio	Number of exceptions	96	98	120	32	40	51
	Failure rate	6.35 %	6.49 %	7.94 %	2.12 %	2.65 %	3.38 %
	Mean VaR overdraft	1.13 %	1.25 %	1.30 %	1.17 %	1.13 %	1.21 %
	LR_{uc} for unconditional coverage	**5.39***	**6.45***	**23.54****	**14.44****	**28.52****	**53.17****
	LR_{ind} for independent exceptions	**139.21****	**134.47****	**180.39****	**65.55****	**89.45****	**132.84****
	LR_{cc} for conditional coverage	**144.60****	**140.91****	**203.93****	**79.98****	**117.97****	**186.01****
Medium portfolio	Number of exceptions	96	113	136	37	41	58
	Failure rate	6.35 %	7.48 %	9.00 %	2.45 %	2.71 %	3.84 %
	Mean VaR overdraft	1.28 %	1.26 %	1.35 %	1.14 %	1.29 %	1.34 %
	LR_{uc} for unconditional coverage	**5.39***	**17.07****	**41.58****	**22.81****	**30.52****	**71.49****
	LR_{ind} for independent exceptions	**126.75***	**167.32****	**218.36****	**70.34****	**97.87****	**175.11****
	LR_{cc} for conditional coverage	**132.14***	**184.39****	**259.94****	**93.15****	**128.40****	**246.60****
Large portfolio	Number of exceptions	101	110	130	38	39	57
	Failure rate	6.68 %	7.28 %	8.60 %	2.51 %	2.58 %	3.77 %
	Mean VaR overdraft	1.16 %	1.19 %	1.29 %	1.07 %	1.24 %	1.23 %
	LR_{uc} for unconditional coverage	**8.20****	**14.58****	**34.30****	**24.66****	**26.56****	**68.76****
	LR_{ind} for independent exceptions	**136.22***	**167.59****	**209.62****	**70.13****	**94.27****	**162.12****
	LR_{cc} for conditional coverage	**144.42****	**182.18****	**243.93****	**94.79****	**120.84****	**230.88****

* Significant at the 5% level.
** Significant at the 1% level.

It can be noted from Table 19 that similarly to the other four models, two of Diagonal sector-beta model's weekly VaR measure series fulfil to a certain degree the independence of exceptions conditions and the corresponding hypothesis isn't rejected at the 1% significance level. This happens for the medium and large portfolio's weekly 95% VaR measures applying 250 days in parameter estimation. In consequence of the failure rates and the exception dependencies during the test period, only the medium portfolio's weekly 95% VaR measures that apply the shortest estimation period length avoid the rejection of the correct conditional coverage hypothesis at the 1% level of significance.

5.3. Summary results

The study's VaR models' accuracy results presented in the previous chapter show that for each studied VaR measure series the failure rates are above the ideal levels of 5% and 1%, indicating that each model produces in general VaR measures, which underestimate true market risk of an equity portfolio. The results are particularly poor for the 99% VaR measures and clearly demonstrate that none of the five VaR models provides valid 99% VaR measures.

The model's inability to produce accurate VaR measures at the 99% confidence level seems to be mainly due to non-normality of the portfolio returns. Indications of non-normality is attained from the observed return distributions' skewness, kurtosis and Anderson-Darling test results presented in Chapter 4.2, and is further supported by the graphical distribution comparisons in appendices B, C and D. In other words, the observed portfolio return distributions are not symmetric and have meaningfully fatter tails than the normal distribution has.[28]

In addition to the shortcomings in correct unconditional coverage, each VaR model has such dependent exceptions during the test period that the correct conditional coverage hypothesis is rejected at least at the 5% significance level. The confronted exceptions are more independent for the weekly than for the daily VaR measures.

Other general finding in the study's five VaR models' accuracy results concerns the estimation period length. The VaR models achieve their best accuracy results when the shortest estimation period length of 250 days is applied in the parameter estimation. Thus, the study's results indicate that application of longer than 250 days estimation period lengths isn't worthwhile in the VaR framework, especially if the moving average method is used for variance and covariance terms and the OLS procedure is used for beta estimates. However, application of shorter estimation period lengths than 250 days might lead to better estimates, since they more effectively adapt to the changing market conditions. The

[28] It has to be though acknowledged that the underlying parameters of the normal distribution can vary so strongly during the test period 1995-2000 that the conclusions drawn from the observed portfolio return distributions doesn't necessarily apply to the conditional distributions.

problem with estimation period lengths below 250 days is that they can't be applied for VaR measures required by BIS (Basle Committee on Banking Supervision, 1996).

In the study, the daily VaR measures have in general more accurate unconditional coverage levels than the corresponding weekly VaR measures. This presumably follows primarily from the fact that the weekly VaR measures expect that the market conditions don't change significantly in five days while the daily VaR measures expect no significant change only in one day. Thus, the weekly VaR measures are more exposed and sensitive to the experienced changes in the general market circumstances.

Even though the failure rates above the ideal levels for the study's all VaR models, the likelihood ratio test results indicate in the previous chapter that this may be due to bad luck for some 95% VaR measures and the respective models might in reality be valid. Thus, in order to facilitate comparisons between the VaR models, the accuracy results are presented comprehensively in this chapter in two tables. One table reports the accuracy results for each VaR model's daily 95% measures and the other for weekly 95% VaR measures. Since the 250 days estimation period length produces best results, only these results are compared and summarised here.

In the next page's Table 20, the comparable test period accuracy results are presented for the daily 95% VaR measures. As can be noticed from the table, Diagonal beta model seems to produce very equal market risk measures to Traditional model. Irrespective of the portfolio size, the correct unconditional coverage hypothesis is accepted for the both models. Nevertheless, due to the strongly clustered exceptions, these models like the other three models produce daily 95% VaR measures that lead to the rejection of the correct conditional coverage hypothesis at the 1% significance level.

On the other hand, Beta model is competitive with Traditional model mainly when the accuracy results of the large portfolio are concerned. For the two smaller portfolios, the firm-specific risks remain two significant and Beta model is not able to produce competitive VaR measures; the correct unconditional coverage hypothesis is though rejected only at the 5% significance level for the medium portfolio.

Table 20.
Test period's 1995-2000 accuracy results for daily 95% VaR measures calculated applying 250 days estimation period length

		VaR model				
		Traditional	Beta	Diagonal beta	Sector-beta	Diagonal sector-beta
Small portfolio	Number of exceptions	86	119	87	117	96
	Failure rate	5.68 %	7.85 %	5.74 %	7.72 %	6.34 %
	Mean VaR overdraft	0.64 %	0.60 %	0.64 %	0.61 %	0.61 %
	LR_{uc} for unconditional coverage	1.40	**22.31****	1.68	**20.42****	**5.27***
	LR_{ind} for independent exceptions	**126.32****	**201.89****	**125.59****	**192.62****	**134.72****
	LR_{cc} for conditional coverage	**127.72****	**224.21****	**127.27****	**213.04****	**139.99****
Medium portfolio	Number of exceptions	83	93	82	100	89
	Failure rate	5.48 %	6.14 %	5.41 %	6.60 %	5.87 %
	Mean VaR overdraft	0.69 %	0.70 %	0.70 %	0.68 %	0.69 %
	LR_{uc} for unconditional coverage	0.71	**3.87***	0.53	**7.46****	2.32
	LR_{ind} for independent exceptions	**125.86****	**139.95****	**122.01****	**152.69****	**130.17****
	LR_{cc} for conditional coverage	**126.57****	**143.82****	**122.53****	**160.15****	**132.49****
Large portfolio	Number of exceptions	81	89	81	90	87
	Failure rate	5.35 %	5.87 %	5.35 %	5.94 %	5.74 %
	Mean VaR overdraft	0.67 %	0.66 %	0.67 %	0.69 %	0.67 %
	LR_{uc} for unconditional coverage	0.37	2.32	0.37	2.67	1.68
	LR_{ind} for independent exceptions	**127.83****	**140.69****	**127.83****	**141.98****	**134.73****
	LR_{cc} for conditional coverage	**128.20****	**143.01****	**128.20****	**144.65****	**136.41****

* **Significant at the 5% level.**
** **Significant at the 1% level.**

Comparing the accuracy results in Table 20, Sector-beta and Diagonal sector-beta model seem to provide in general weaker daily 95% VaR measures than the respective VaR models that apply position mapping onto the market index. Furthermore, taking into account the larger number of parameters needed for VaR models that are based on mapping equities onto the sector indices, application of mapping onto the market seems to be more preferable in VaR calculations. This indication can be obtained also from the next page's Table 21, which reports the accuracy results for each model's weekly VaR measures that apply the study's shortest estimation period length of 250 days.

As stated earlier in the study, the only significant distinction between the study's daily and weekly VaR measures is the validity of the exception independence hypothesis. In Table 21, the hypothesis isn't rejected at the 1% significance level for each VaR measure series, which is in contrast to the daily VaR measure results summarised in Table 20. Furthermore, some weekly VaR measure series result in acceptance of the correct conditional coverage hypothesis at the 99% confidence level.

Table 21.
Test period's 1995-2000 accuracy results for weekly 95% VaR measures calculated applying 250 days estimation period length

		VaR model				
		Traditional	Beta	Diagonal beta	Sector-beta	Diagonal sector-beta
Small portfolio	Number of exceptions	88	113	91	114	96
	Failure rate	5.82 %	7.48 %	6.02 %	7.54 %	6.35 %
	Mean VaR overdraft	1.15 %	1.18 %	1.13 %	1.18 %	1.13 %
	LR_{uc} for unconditional coverage	2.06	**17.07****	3.13	**17.94****	**5.39***
	LR_{ind} for independent exceptions	**117.81***	**167.86****	**124.92***	**167.60****	**139.21****
	LR_{cc} for conditional coverage	**119.86***	**184.93****	**128.05****	**185.53****	**144.60****
Medium portfolio	Number of exceptions	86	104	86	110	96
	Failure rate	5.69 %	6.88 %	5.69 %	7.28 %	6.35 %
	Mean VaR overdraft	1.31 %	1.25 %	1.31 %	1.25 %	1.28 %
	LR_{uc} for unconditional coverage	1.46	**10.14****	1.46	**14.58****	**5.39***
	LR_{ind} for independent exceptions	**110.00***	**132.68***	**110.00***	**156.85****	**126.75***
	LR_{cc} for conditional coverage	**111.46***	**142.82****	**111.46***	**171.44****	**132.14***
Large portfolio	Number of exceptions	92	97	91	105	101
	Failure rate	6.09 %	6.42 %	6.02 %	6.95 %	6.68 %
	Mean VaR overdraft	1.16 %	1.21 %	1.17 %	1.19 %	1.16 %
	LR_{uc} for unconditional coverage	3.54	**5.91***	3.13	**10.83****	**8.20****
	LR_{ind} for independent exceptions	**121.41***	**122.59***	**118.07***	**141.06***	**136.22***
	LR_{cc} for conditional coverage	**124.95***	**128.50***	**121.20***	**151.89****	**144.42****

* Significant at the 5% level.
** Significant at the 1% level.

Overall, the weekly VaR measure accuracy comparisons imply similar findings to the daily VaR measure comparisons. In other words, Traditional and Diagonal beta model's VaR measures show best accuracy among the five models. Again, Beta model improves its competitiveness as the portfolio size increases and provides rather accurate weekly 95% VaR measures for the large portfolio. Furthermore, also the comparable results of Table 21 indicate that mapping onto the market is preferable to mapping onto the sectors when simplifying covariance matrix calculations in variance-covariance VaR framework. It should be however noted that most of this accuracy difference results from the early part of the test period, as observed in the standard deviation comparisons in Chapter 5.1. Thus, when only the latter three years of the test period are considered, the two models applying position mapping onto the sector indices produce more competitive VaR measures. Nevertheless, they still fail to be superior to the study's two models that apply equity mapping onto the market, and the accuracy results concerning the latter part of the test period are not reported separately in the study. Finally, the mean VaR overdrafts presented in the tables above seem to be more portfolio-dependent than VaR model-dependent, and they thus shouldn't have impact on how the study's different VaR models are preferred to each other.

6. Conclusions

This study examines how accurate daily and weekly 95% and 99% VaR measures do the five different VaR models produce for equity portfolios. Each of the five VaR models applies the common variance-covariance approach for VaR measures. One of the VaR models examined is the traditional variance-covariance model that calculates the necessary covariance matrix in a standard way by using portfolio's equities' historical return data for variances and covariances. In the study, of special interest are however four different VaR models that map equity positions onto the market index and sector indices and use the beta-corrected variances and covariances of the indices in deriving the equity portfolio's covariance matrix.

The VaR models' accuracy abilities are investigated by backtesting the VaR measures calculated with the models. Backtesting period covers years 1995-2000. From the liquid S&P 500 constituent equities, three different portfolios with 30, 50 and 100 equities are formed in a random manner for the study. Parameter estimation period lengths of past 250, 500 and 1000 days are used during the test period for the VaR models' VaR measures.

Backtesting of the VaR models reveals that none of them seems to provide highly accurate VaR measures. Clearly inaccurate and inappropriate are the models' both daily as well as weekly 99% VaR measures, since they confront much higher number of exceeded VaR measures, that is exceptions, than expected. This is a fairly usual conclusion made also in the past studies for the variance-covariance approach's 99% VaR measures. As commonly argued in the past research and further strengthened through this study's examination on the properties of the equity portfolio returns, the weakness of the 99% VaR measures results presumably mainly from equity portfolio return distributions' fatter tails than implied by the normal distribution. So, as suggested by Hendricks (1996), application of the *t* distribution with between four and six degrees of freedom, instead of the normal distribution, may well provide more accurate coverage for the 99% VaR measures of the variance-covariance approach.

Although the VaR models studied can produce daily and weekly 95% VaR measures that don't seem to be exceeded statistically significantly more or less than 5% of times, the VaR models are missing an important property that weakens their attractiveness. Particularly, the VaR models studied appear to be rather incapable of encountering VaR measure exceptions independently from each other. In other words, the exceptions occur in clusters, which indicates that the models don't adapt effectively enough to the changing market conditions; the exceptions seem to be more clustered for the daily than for the weekly VaR measures. The past VaR accuracy studies very seldom highlight the problem associated with clustered exceptions, which is supposedly due to the fact that effective tests for the exception independence have been introduced only very recently.

The drawback of dependent exceptions could be perhaps lessened by applying variance and covariance estimation methods, such as the exponentially weighted moving average (EWMA) methods, that more powerfully take into account the prevailing market conditions than the moving average (MA) methods applied in the study that weight each past return observation equally. However, as discussed in the study, application of such methods makes the time aggregation more difficult, because the square root of time rule is not appropriate for these kinds of daily variance and covariance estimates.

Comparison of the study's VaR models presents interesting issues. One such issue is that from the three different estimation period lengths applied in the study, the shortest length of past 250 days produces during the test period fairly distinctly the most accurate VaR measures in each of the five VaR models. The similar finding is made also in the study of Johansson, Seiler and Tjarnberg (1999), while the findings of Jackson, Maude and Perraudin (1997) on contrary suggest that the longer the estimation period length the more accurate the VaR measure.

In general, the study's two beta mapping-based VaR models that take into account also the firm-specific risks in addition to the general market or sector-specific risks seem to provide very equal VaR measures compared to the traditional variance-covariance VaR model; this holds especially when equity positions are mapped onto the market index. The high equality results from their ability to produce fairly identical portfolio return standard deviations with the traditional model irrespective of the portfolio size. The results of the study indicate that consideration of the firm-specific risks in beta mapping-based VaR models improves the VaR measure accuracy quite meaningfully if the portfolio is relatively undiversified.

When the two plain beta mapping-based VaR models that neglect the firm-specific risks are considered, their competitiveness appears to be rather dependent on the portfolio size. In accordance with the findings of Johansson, Seiler and Tjarnberg (1999), the VaR accuracy results of this study indicate that plain beta VaR models lack relevant information if only some dozens of equities are included in the portfolio. When the portfolio is more diversified, such as the study's largest portfolio including 100 equities, the plain beta VaR models seem to output relatively competent equity portfolio VaR measures with the traditional variance-covariance VaR model. Due to the data availability problems, the study's portfolios include at highest equities only from eight different S&P sectors out of eleven. Therefore, inclusion of the remaining three sectors would supposedly make the portfolios more effectively diversified, and thus increase accuracy of the plain beta mapping-based VaR models.

Based on the results achieved through the backtesting period 1995-2000, the study's two VaR models that apply equity mapping onto the market index seem to be preferable to the respective two VaR models that apply mapping onto the sector indices. The differences between the VaR models applying different mapping approaches are however far from being stable during the test period, and the VaR models mapping onto the sectors advance as gone towards the end of the test period. Thus, it may be worthwhile to further research whether the VaR models that are based on mapping equities onto the sector indices have permanently increased their competitiveness, and perhaps provide nowadays more accurate market risk measures than the VaR models mapping onto the market index. Also, if only a couple of sectors are represented in an equity portfolio, it might be more appropriate to map positions onto the sectors rather than onto the market, since the potentially too low correlations between sector indices don't so significantly affect on the portfolio return estimate of an VaR model applying mapping onto the sectors.

References

Alexander, C. O. and C. T. Leigh (1997). On the Covariance Matrices Used in Value at Risk Models. *The Journal of Derivatives* Vol. 4, No. 3, 50-62.

Basle Committee on Banking Supervision (1996). Overview of the Amendment to the Capital Accord to Incorporate Market Risks.

Campbell, John Y., Andrew W. Lo and A. Craig MacKinlay (1997). *The Econometrics of Financial Markets.* Princeton University Press.

Christoffersen, Peter F. (1998). Evaluating Interval Forecasts. *International Economic Review* Vol. 39, No. 4, 841-862.

D'Agostino, R. B. and M. A. Stephens (1986). *Goodness of Fit Techniques.* Marcel Dekker.

Dowd, Kevin (1998). *Beyond Value at Risk: The New Science of Risk Management.* John Wiley & Sons.

Fleming, Michael J. and Joseph G. Nellis (1994). *Principles of Applied Statistics.* Routledge.

Haas, Marcus (2001). New Methods in Backtesting. *Financial Engineering Research center caesar,* working paper.

Hendricks, Darryll (1996). Evaluation of Value-at-Risk Models Using Historical Data. *Federal Reserve Bank of New York Economic Policy Review* Vol. 2, No. 1, 39-69.

Jackson, Patricia, David J. Maude and William Perraudin (1997). Bank Capital And Value At Risk. *The Journal of Derivatives* Vol. 4, No. 3, 73-90.

Johansson, Frederik, Michael J. Seiler and Mikael Tjarnberg (1999). Measuring downside portfolio risk. *Journal of Portfolio Management* Vol. 26, No. 1, 96-107.

Jorion, Philippe (1996). *Value at Risk: The New Benchmark for Controlling Market Risk.* Irwin.

Jorion, Philippe (2000). *Value at Risk: The New Benchmark for Managing Financial Risk.* McGraw-Hill.

J.P. Morgan/Reuters (1996). *RiskMetrics™-Technical Document.* Morgan Guaranty Trust Company.

Kupiec, Paul H. (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. *The Journal of Derivatives* Vol. 3, No. 2, 73-84.

Linsmeier, Thomas J. and Neil D. Pearson (2000). Value at Risk. *Financial Analysts Journal* Vol. 56, No. 2, 47-67.

Mahoney, James M (1995). Empirical-based Versus Model-based Approaches to Value-at-Risk. *Federal Reserve Bank of New York*, working paper.

Manfredo, Mark R. and Raymond M. Leuthold (2001). Market Risk and the Cattle Feeding Margin: An Application of Value-at-Risk. *Arizona State University East*, working paper.

McNeil, Alexander J. and Rüdiger Frey (2000). Estimation of tail-related risk measures for heteroskedastic financial time series: An extreme value approach. *Journal of Empirical Finance* Vol. 7, 271-300.

Polasek, Wolfgang and Momtchil Pojarliev (2000). VaR Evaluations Based on Volatility Forecasts of GARCH Models. *University of Basel*, working paper.

S&P 500 Index Methodology (2001). www.spglobal.com/5sec3.pdf.

Sarma, Mandira (2001). Testing EVT-based VaR measures. *Indira Gandhi Institute of Development Research*, working paper.

Sharpe, William F. (2000). *Portfolio Theory and Capital Markets.* McGraw-Hill.

Simon, Carl P. and Lawrence Blume (1994). *Mathematics for Economists.* W. W. Norton & Company.

Sinclair, C. D. and B. D. Spurr (1988). Approximations to the Distribution Function of the Anderson-Darling Test Statistic. *Journal of the American Statistical Association* Vol. 83, No. 404, 1990-1991.

Wong, Michael Chak Sham, Wai Yan Cheng and Clement Yuk Ping Wong (2001). Market Risk Management of Banks: Implications from the Accuracy of Value-at-Risk Forecasts. *City University of Hong Kong,* working paper.

Appendix A. Comparison of representative mapping approaches

By applying the same notations as in the main text, it's shown in this appendix where do the differences between the portfolio variances arise from, when individual equity positions are mapped onto the market index vs. onto the sector indices. The comparison is for simplicity made using a portfolio that consists of only two only equities, which represent two separate sectors. Furthermore, the approaches compared here don't take into the firm-specific risks of individual equities. Inclusion of such risks should generally increase competitiveness of mapping onto market index over mapping onto sector indices, as explained at the end of the appendix.

In the following, a subtraction between portfolio variances calculated through Formula (15) and through Formula (21), excluding the latter part that represents firm-specific risks, is simplified in order to find where the variance difference effectively results from.

$$\begin{bmatrix} w_1 & w_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \sigma_m^2 - \begin{bmatrix} w_{S1}\beta_{S1} & w_{S2}\beta_{S2} \end{bmatrix} \begin{bmatrix} \sigma_{S1}^2 & \sigma_{S1S2} \\ \sigma_{S2S1} & \sigma_{S2}^2 \end{bmatrix} \begin{bmatrix} w_{S1}\beta_{S1} \\ w_{S2}\beta_{S2} \end{bmatrix}$$

$$= w_1^2\beta_1^2\sigma_m^2 + w_2^2\beta_2^2\sigma_m^2 + 2w_1w_2\beta_1\beta_2\sigma_m^2 - w_{S1}^2\beta_{S1}^2\sigma_{S1}^2 - w_{S2}^2\beta_{S2}^2\sigma_{S2}^2 - 2w_{S1}w_{S2}\beta_{S1}\beta_{S2}\sigma_{S1S2}$$

Since individual equity weights are here same as the sector weights, the subtraction can be simplified to the following form

$$w_1^2(\beta_1^2\sigma_m^2 - \beta_{S1}^2\sigma_{S1}^2) + w_2^2(\beta_2^2\sigma_m^2 - \beta_{S2}^2\sigma_{S2}^2) + 2w_1w_2(\beta_1\beta_2\sigma_m^2 - \beta_{S1}\beta_{S2}\sigma_{S1S2}).$$

Because betas are estimated applying Formula (42) and are thus derived from correlation coefficients and standard deviations, the subtraction gets the following form

$$w_1^2\left(\rho_{1m}^2\frac{\sigma_1^2}{\sigma_m^2}\sigma_m^2 - \rho_{1S1}^2\frac{\sigma_1^2}{\sigma_{S1}^2}\sigma_{S1}^2\right) + w_2^2\left(\rho_{2m}^2\frac{\sigma_2^2}{\sigma_m^2}\sigma_m^2 - \rho_{2S2}^2\frac{\sigma_2^2}{\sigma_{S2}^2}\sigma_{S2}^2\right)$$

$$+ 2w_1w_2\left(\rho_{1m}\frac{\sigma_1}{\sigma_m}\rho_{2m}\frac{\sigma_2}{\sigma_m}\sigma_m^2 - \rho_{1S1}\frac{\sigma_1}{\sigma_{S1}}\rho_{2S2}\frac{\sigma_2}{\sigma_{S2}}\rho_{S1S2}\sigma_{S1}\sigma_{S2}\right).$$

Finally, the subtraction can be simplified to the form shown below, from which the actual determinants of the portfolio variance difference can be easily observed

$$w_1^2\sigma_1^2(\rho_{1m}^2-\rho_{1S1}^2)+w_2^2\sigma_2^2(\rho_{2m}^2-\rho_{2S2}^2)+2w_1w_2\sigma_1\sigma_2(\rho_{1m}\rho_{2m}-\rho_{1S1}\rho_{2S2}\rho_{S1S2}).$$

As can be noted from above, the portfolio variance difference between a VaR model mapping equity positions onto the market and a model that maps onto the sectors is effectively due to the correlation conditions. In other words, since it's expected that a sector index can explain the return of an equity in the particular sector better than the market index, and thus its correlation with the equity is larger than the corresponding correlation between the equity and the market index, the first two terms above, representing individual variance differences, are supposedly in favour of mapping onto sector indices. However, the last, third term above representing differences in covariances can well be in favour of mapping onto market index, if the correlation between sector indices ρ_{S1S2} is low enough.

It's important to point out that when the portfolio gets larger and the number of sectors represented increases, the number of covariance differences grows exponentially, while the number of variance differences increases only linearly. This means that the relative significance of the correlations between sector indices increases, when the number of sectors rises in the portfolio, and if these correlations are rather low, the individual equities must have much higher correlations with the sector indices than with market, so that mapping onto the sector indices would produce higher portfolio variances and VaR measures than mapping onto the market index.

The previously described comparison doesn't consider differences in firm-specific risks. As equities correlate presumably more effectively with the sector indices than with the market index, the firm-specific risks, which are defined to be risks that cannot be explained by equity's correlation with the sector or market index, should be in general higher when mapping is applied onto the market index. Therefore, the inclusion of firm-specific risks in VaR models that apply mapping procedures should benefit more when mapping is made onto the market index rather than onto the sector indices.

Appendix B. Small portfolio's return figures

Figure B1.
Small portfolio's observed daily return distribution during 1995-2000 vs. normal distribution (μ =0.09%, σ =1.04%)

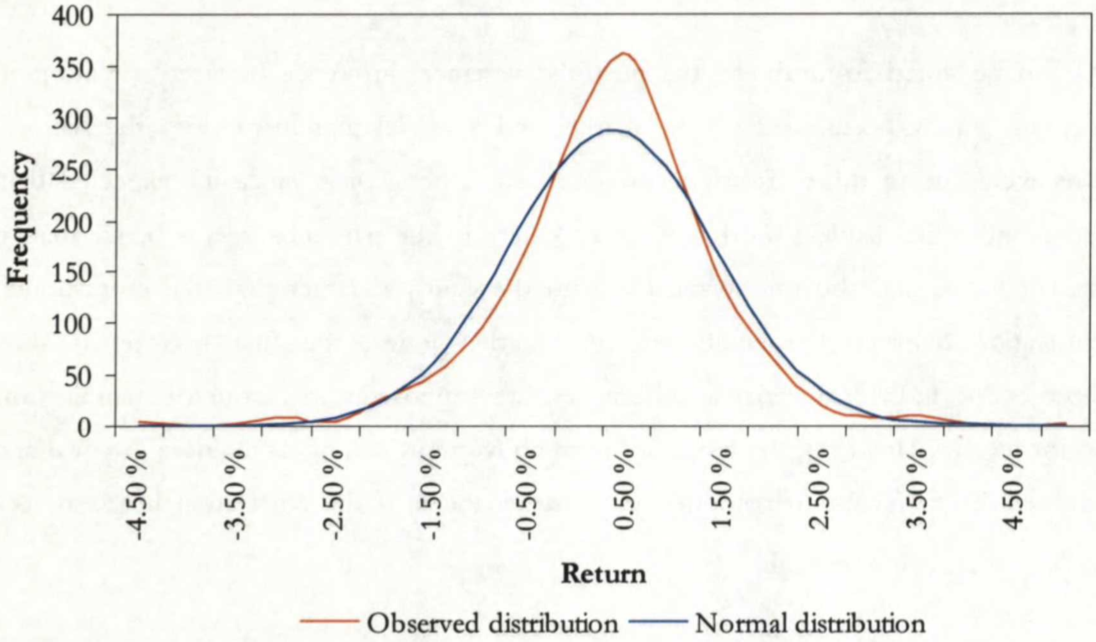
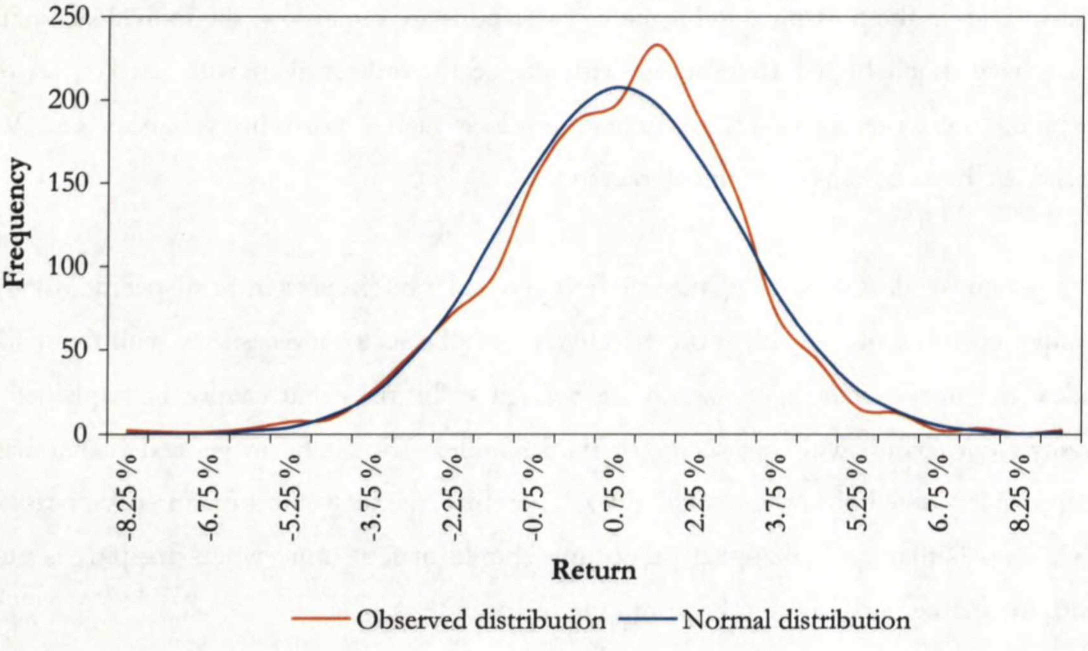


Figure B2.
Small portfolio's observed weekly return distribution during 1995-2000 vs. normal distribution (μ =0.44%, σ =2.17%)



Appendix C. Medium portfolio's return figures

Figure C1.
Medium portfolio's observed daily return distribution during 1995-2000 vs. normal distribution (μ =0.08, σ =1.08%)

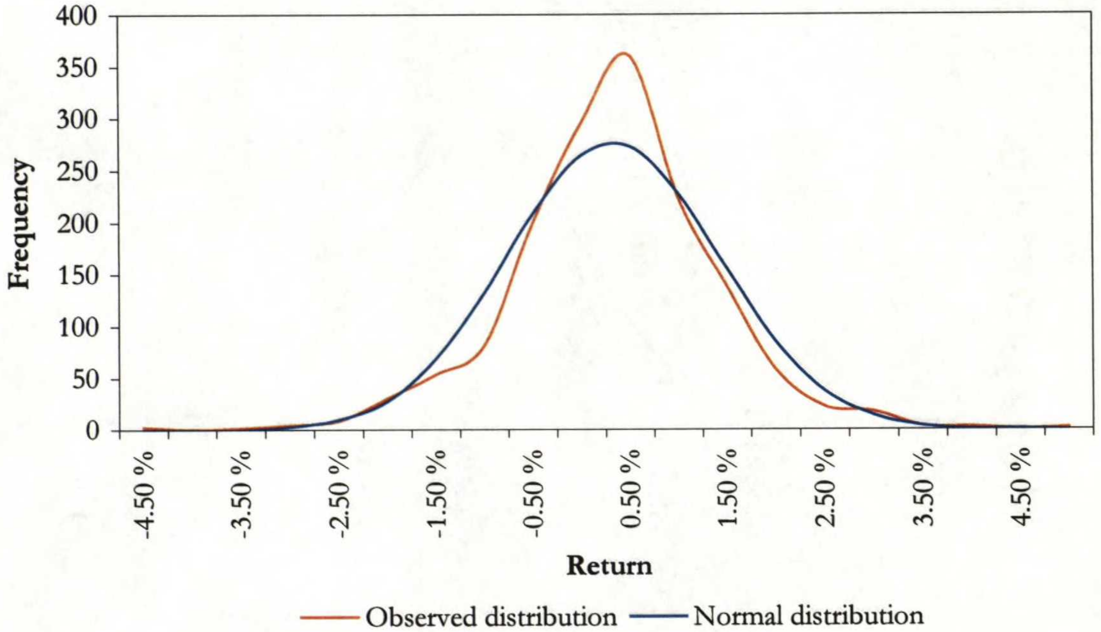
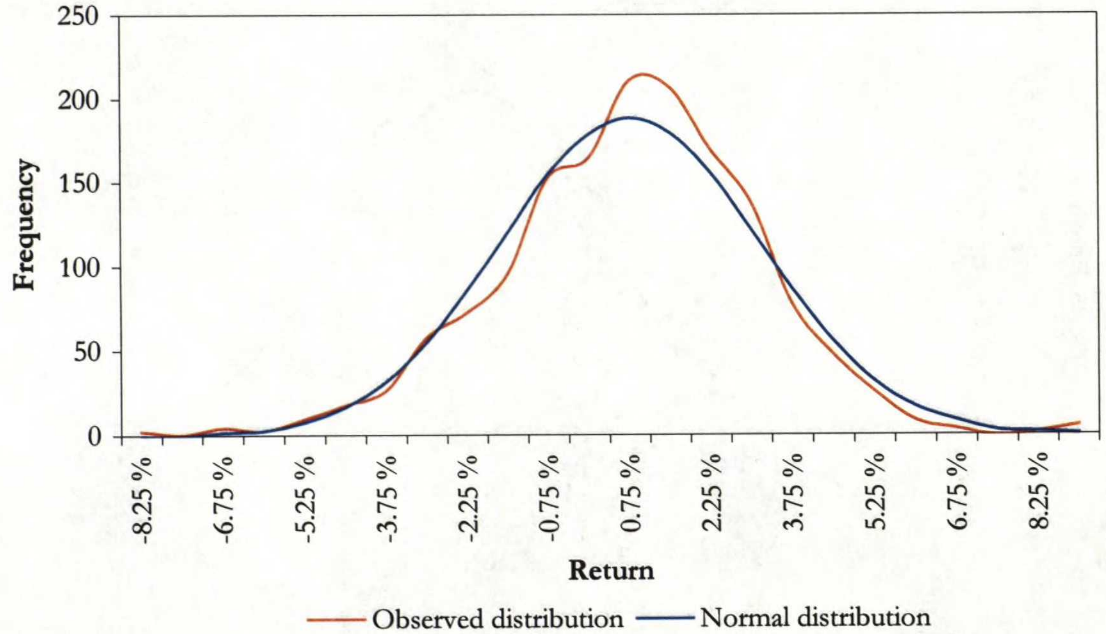


Figure C2.
Medium portfolio's observed weekly return distribution during 1995-2000 vs. normal distribution (μ =0.39%, σ =2.39%)



Appendix D. Large portfolio's return figures

Figure D1.
Large portfolio's observed daily return distribution during 1995-2000 vs. normal distribution (μ =0.09%, σ =1.06%)

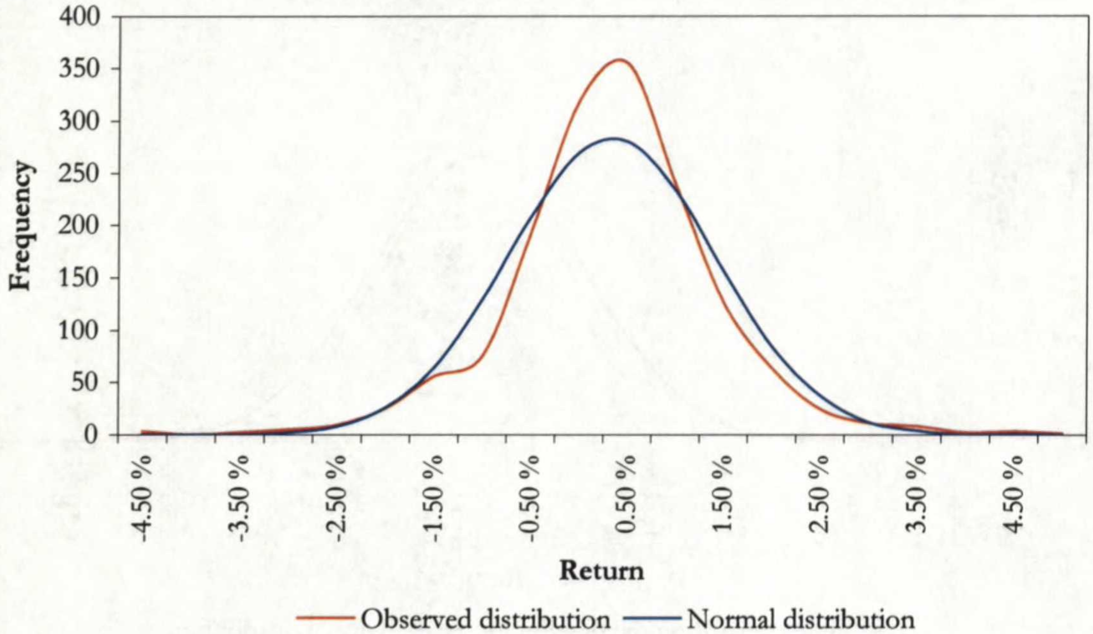
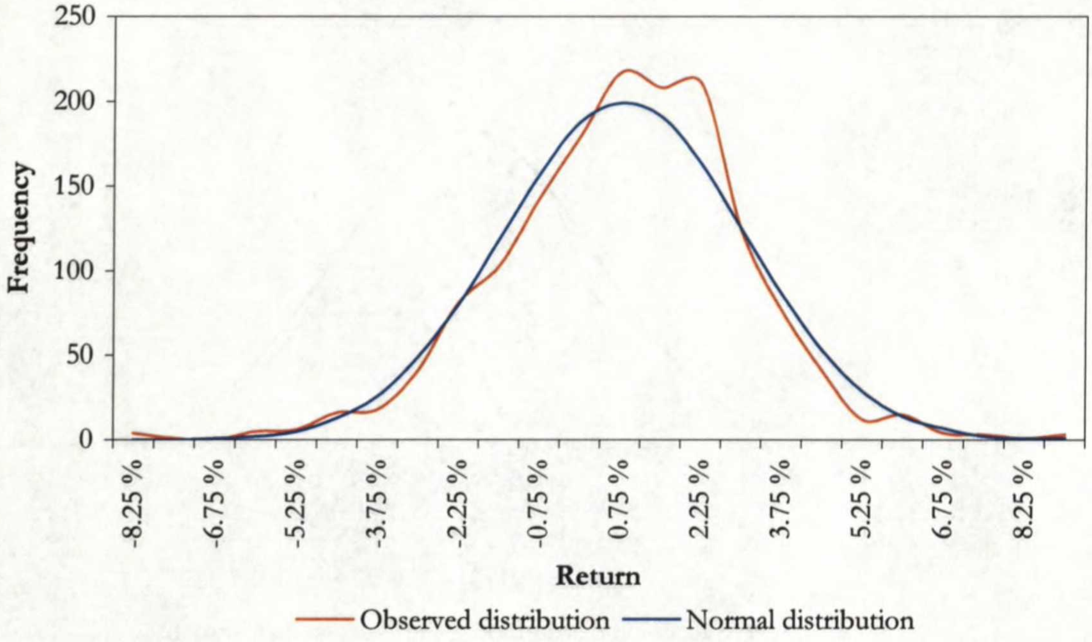


Figure D2.
Large portfolio's observed weekly return distribution during 1995-2000 vs. normal distribution (μ =0.43%, σ =2.25%)



Appendix E. Small portfolio's supplementary STD figures

Figure E1.
Daily return standard deviations used in different VaR models during test period (small portfolio, estimation period length 500 days)

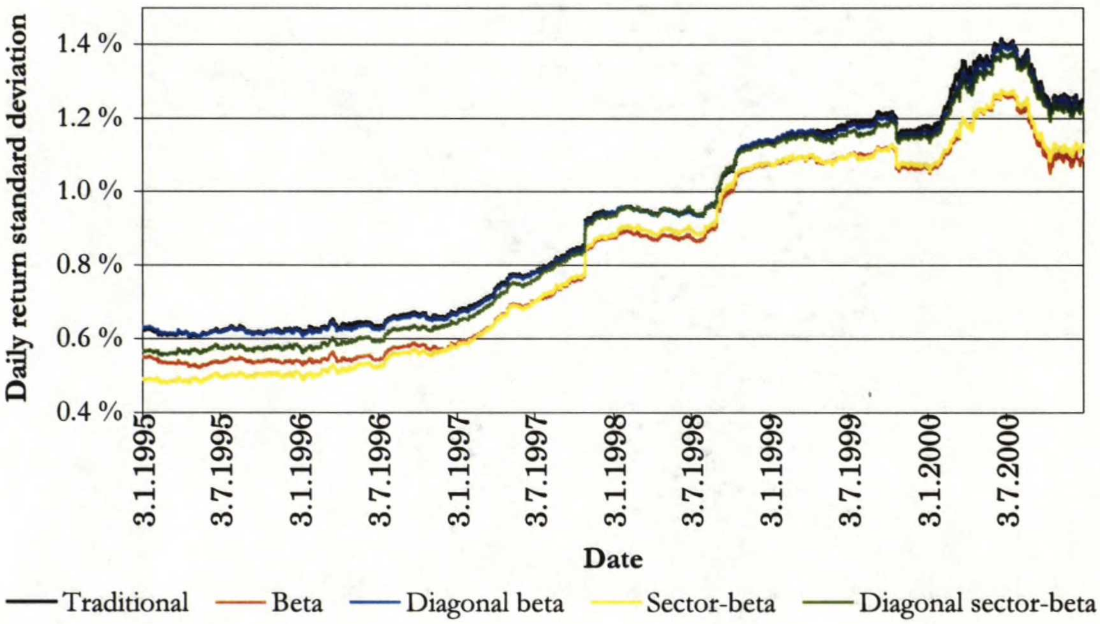
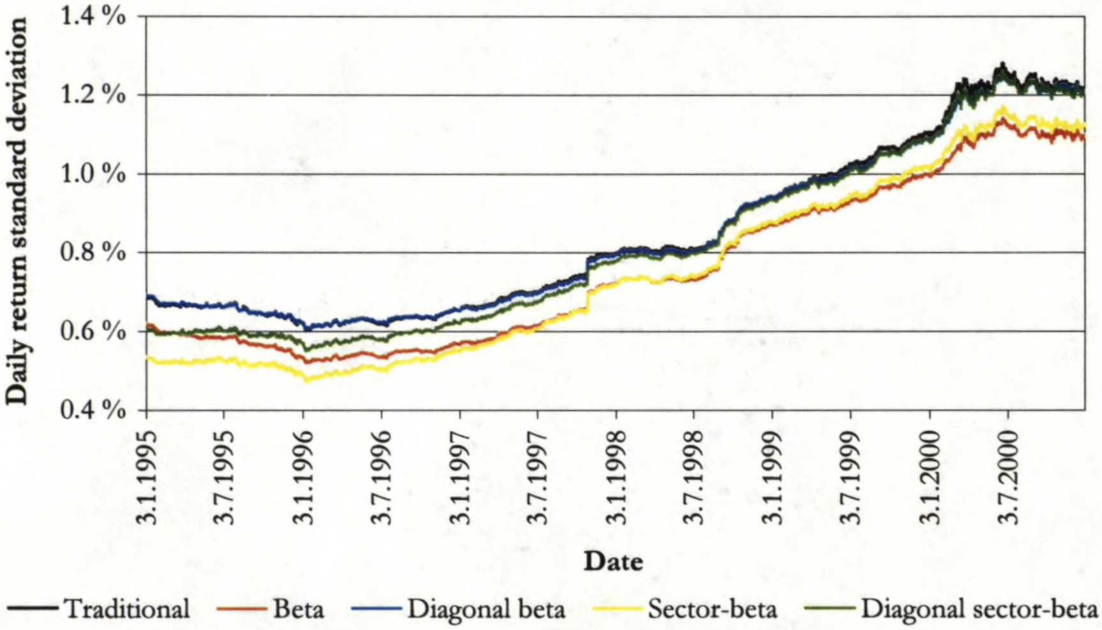


Figure E2.
Daily return standard deviations used in different VaR models during test period (small portfolio, estimation period length 1000 days)



Appendix F. Medium portfolio's supplementary STD figures

Figure F1.
Daily return standard deviations used in different VaR models during test period (medium portfolio, estimation period length 500 days)

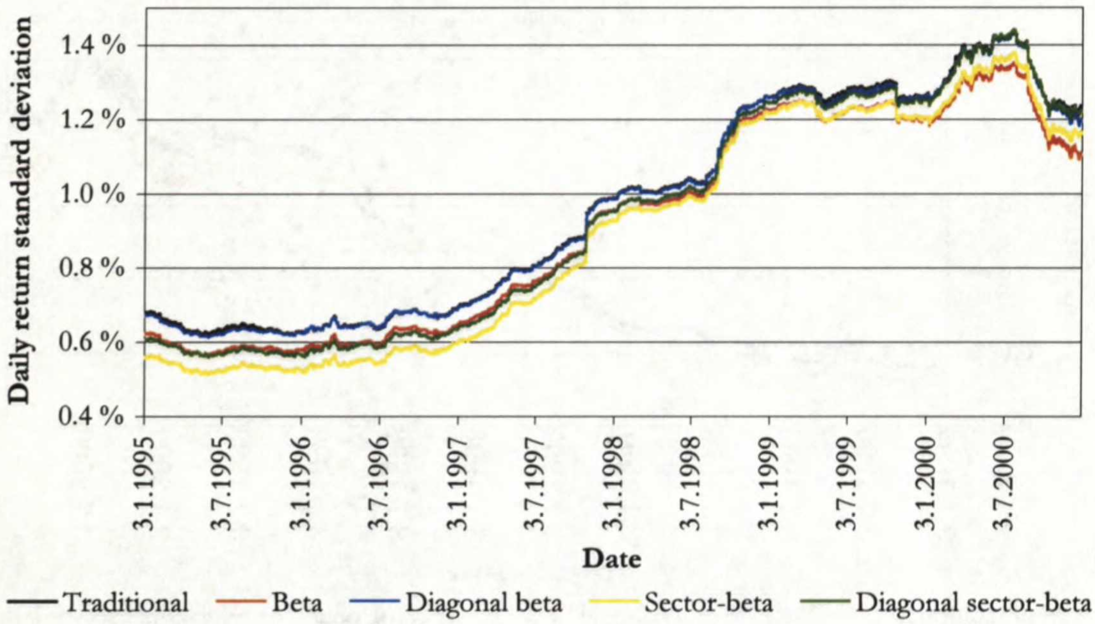
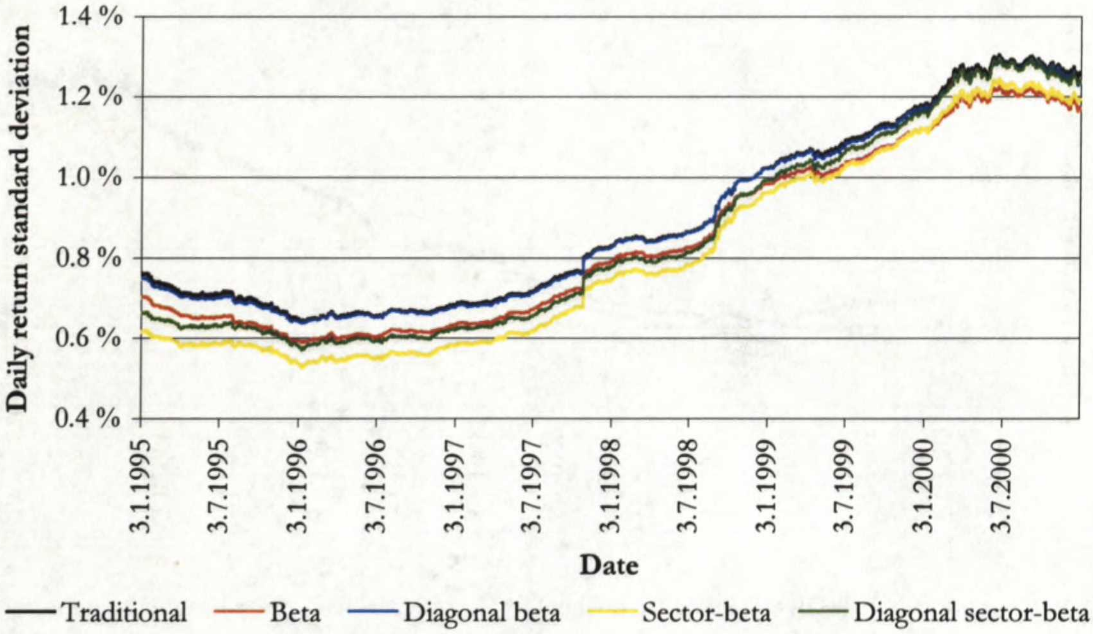


Figure F2.
Daily return standard deviations used in different VaR models during test period (medium portfolio, estimation period length 1000 days)



Appendix G. Large portfolio's supplementary STD figures

Figure G1.
Daily return standard deviations used in different VaR models during test period (large portfolio, estimation period length 500 days)

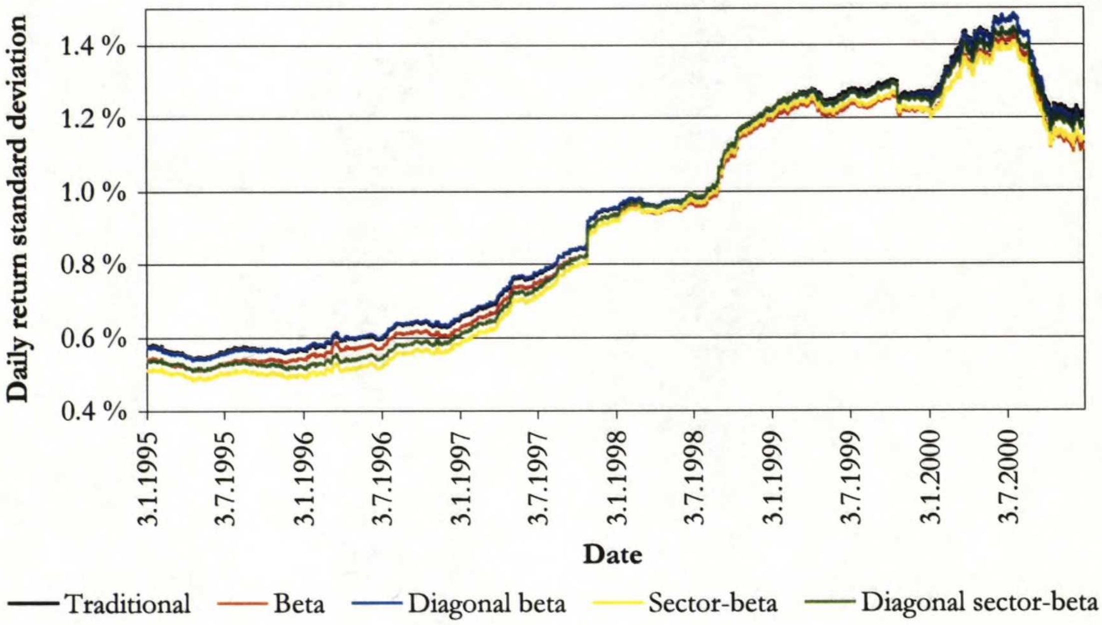


Figure G2.
Daily return standard deviations used in different VaR models during test period (large portfolio, estimation period length 1000 days)